TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2157

Travel Behavior 2010

VOLUME 2

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH BOARD

OF THE NATIONAL ACADEMIES

TRANSPORTATION RESEARCH RECORDS, which are published throughout the year, consist of collections of papers on specific transportation modes and subject areas. Each Record is classified according to the subscriber category or categories covered by the papers published in that volume. The views expressed in papers published in the Transportation Research Record series are those of the authors and do not necessarily reflect the views of the peer review committee(s), the Transportation Research Board, the National Research Council, or the sponsors of TRB activities. The Transportation Research Board does not endorse products or manufacturers; trade and manufacturers' names may appear in a Record paper only if they are considered essential.

PEER REVIEW OF PAPERS: All papers published in the Transportation Research Record series have been reviewed and accepted for publication through the Transportation Research Board's peer review process established according to procedures approved by the Governing Board of the National Research Council. Papers are refereed by the TRB standing committees identified on page ii of each Record. Reviewers are selected from among committee members and other outside experts. The Transportation Research Board requires a minimum of three reviews; a decision is based on reviewer comments and resultant author revision. For details about the peer review process, see the information on the inside back cover.

THE TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD, comprising a cross section of transportation disciplines and with equal representation from the academic and practitioner communities, assures that the quality of the *Transportation Research Record: Journal of the Transportation Research Board* and the TRB paper peer review process is consistent with that of a peer-reviewed scientific journal. Members from the academic community make decisions on granting tenure; all academic and practitioner members have authored papers published in peer-reviewed journals and all have participated in the TRB peer review process.

TRANSPORTATION RESEARCH BOARD PUBLICATIONS may be ordered directly from the TRB Business Office, through the Internet at www.TRB.org, or by annual subscription through organizational or individual affiliation with TRB. Affiliates and library subscribers are eligible for substantial discounts. For further information, contact the Transportation Research Board Business Office, 500 Fifth Street, NW, Washington, DC 20001 (telephone 202-334-3213; fax 202-334-2519; or e-mail TRBsales@nas.edu).

TRANSPORTATION RESEARCH RECORD PAPERS ONLINE: The TRR Journal Online website provides electronic access to the full text of papers that have been published in the Transportation Research Record series since 1996. The site is updated as new Record papers become available. To search abstracts and to find subscription and pricing information, go to www.TRB.org/TRROnline.

TRANSPORTATION RESEARCH BOARD 2010 EXECUTIVE COMMITTEE*

Chair: Michael R. Morris, Director of Transportation, North Central Texas Council of Governments, Arlington
Vice Chair: Neil J. Pedersen, Administrator, Maryland State Highway Administration, Baltimore
Executive Director: Robert E. Skinner, Jr., Transportation Research Board

J. Barry Barker, Executive Director, Transit Authority of River City, Louisville, Kentucky
Allen D. Biehler, Secretary, Pennsylvania Department of Transportation, Harrisburg
Larry L. Brown, Sr., Executive Director, Mississippi Department of Transportation, Jackson
Deborah H. Butler, Executive Vice President, Planning, and CIO, Norfolk Southern Corporation, Norfolk, Virginia
William A. V. Clark, Professor, Department of Geography, University of California, Los Angeles
Eugene A. Conti, Jr., Secretary of Transportation, North Carolina Department of Transportation, Raleigh
Nicholas J. Garber, Henry L. Kinnier Professor, Department of Civil Engineering, and Director, Center for Transportation Studies, University of Virginia, Charlottesville
Jeffrey W. Hamiel, Executive Director, Metropolitan Airports Commission, Minneapolis, Minnesota
Paula J. Hammond, Secretary, Washington State Department of Transportation, Olympia
Edward A. (Ned) Helme, President, Center for Clean Air Policy, Washington, D.C.
Adib K. Kanafani, Cahill Professor of Civil Engineering, University of California, Berkeley (Past Chair, 2009)
Susan Martinovich, Director, Nevada Department of Transportation, Carson City
Debra L. Miller, Secretary, Kansas Department of Transportation, Topeka (Past Chair, 2008)
Sandra Rosenbloom, Professor of Planning, University of Arizona, Tucson
Tracy L. Rosser, Vice President, Corporate Traffic, Wal-Mart Stores, Inc., Mandeville, Louisiana
Steven T. Scalzo, Chief Operating Officer, Marine Resources Group, Seattle, Washington
Henry G. (Gerry) Schwartz, Jr., Chairman (retired), Jacobs/Sverdrup Civil, Inc., St. Louis, Missouri
Beverly A. Scott, General Manager and Chief Executive Officer, Metropolitan Atlanta Rapid Transit Authority, Atlanta, Georgia
David Seltzer, Principal, Mercator Advisors LLC, Philadelphia, Pennsylvania
Daniel Sperling, Professor of Civil Engineering and Environmental Science and Policy; Director, Institute of Transportation Studies; and Interim Director, Energy Efficiency Center, University of California, Davis
Kirk T. Steudle, Director, Michigan Department of Transportation, Lansing
Douglas W. Stotlar, President and Chief Executive Officer, Con-Way, Inc., Ann Arbor, Michigan
C. Michael Walton, Ernest H. Cockrell Centennial Chair in Engineering, University of Texas, Austin (Past Chair, 1991)

Peter H. Appel, Administrator, Research and Innovative Technology Administration, U.S. Department of Transportation (ex officio)
J. Randolph Babbitt, Administrator, Federal Aviation Administration, U.S. Department of Transportation (ex officio)
Rebecca M. Brewster, President and COO, American Transportation Research Institute, Smyrna, Georgia (ex officio)
George Bugliarello, President Emeritus and University Professor, Polytechnic Institute of New York University, Brooklyn; Foreign Secretary, National Academy of Engineering, Washington, D.C. (ex officio)
Anne S. Ferro, Administrator, Federal Motor Carrier Safety Administration, U.S. Department of Transportation (ex officio)
LeRoy Gishi, Chief, Division of Transportation, Bureau of Indian Affairs, U.S. Department of the Interior, Washington, D.C. (ex officio)
Edward R. Hamberger, President and CEO, Association of American Railroads, Washington, D.C. (ex officio)
John C. Horsley, Executive Director, American Association of State Highway and Transportation Officials, Washington, D.C. (ex officio)
David T. Matsuda, Deputy Administrator, Maritime Administration, U.S. Department of Transportation (ex officio)
Victor M. Mendez, Administrator, Federal Highway Administration, U.S. Department of Transportation (ex officio)
William W. Millar, President, American Public Transportation Association, Washington, D.C. (ex officio) (Past Chair, 1992)
Robert J. Papp (Adm., U.S. Coast Guard), Commandant, U.S. Coast Guard, U.S. Department of Homeland Security (ex officio)
Cynthia L. Quarterman, Administrator, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation (ex officio)
Peter M. Rogoff, Administrator, Federal Transit Administration, U.S. Department of Transportation (ex officio)
David L. Strickland, Administrator, National Highway Traffic Safety Administration, U.S. Department of Transportation (ex officio)
Joseph C. Szabo, Administrator, Federal Railroad Administration, U.S. Department of Transportation (ex officio)
Polly Trottenberg, Assistant Secretary for Transportation Policy, U.S. Department of Transportation (ex officio)
Robert L. Van Antwerp (Lt. General, U.S. Army), Chief of Engineers and Commanding General, U.S. Army Corps of Engineers, Washington, D.C. (ex officio)

* Membership as of August 2010.

TRB

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2157

Travel Behavior 2010

VOLUME 2

A Peer-Reviewed Publication

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES

Washington, D.C.
2010

www.TRB.org

Transportation Research Record 2157
ISSN 0361-1981
ISBN 978-0-309-14283-0

Subscriber Categories
Highways; public transportation; pedestrians and bicyclists; planning and forecasting; passenger transportation; economics; environment; data and information technology

Printed in the United States of America

TRANSPORTATION RESEARCH RECORD PUBLICATION BOARD
C. Michael Walton, PhD, PE, Ernest H. Cockrell Centennial Chair in Engineering and Professor of Civil Engineering, University of Texas, Austin (Cochair)
Mary Lynn Tischer, PhD, Director, Office of Transportation Policy Studies, Federal Highway Administration, Washington, D.C. (Cochair)
Daniel Brand, SB, SM, PE, Consultant, Lyme, New Hampshire
Mary R. Brooks, BOT, MBA, PhD, William A. Black Chair of Commerce, Dalhousie University, Halifax, Nova Scotia, Canada
Charles E. Howard, Jr., MCRP, Director, Transportation Planning, Puget Sound Regional Council, Seattle, Washington
Thomas J. Kazmierowski, PEng, Manager, Materials Engineering and Research Office, Ontario Ministry of Transportation, Toronto, Canada
Michael D. Meyer, PhD, PE, Frederick R. Dickerson Professor, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta
Sandra Rosenbloom, MA, PhD, Professor of Planning, University of Arizona, Tucson
Kumares C. Sinha, PhD, PE, Olson Distinguished Professor, School of Civil Engineering, Purdue University, West Lafayette, Indiana
L. David Suits, MSCE, Executive Director, North American Geosynthetics Society, Albany, New York

Peer Review of Transportation Research Record 2157

PLANNING AND ENVIRONMENT GROUP
Katherine F. Turnbull, Texas Transportation Institute (Chair)

Travel Analysis Methods Section
Ram M. Pendyala, Arizona State University (Chair), Barbara J. Arens, Kevin J. Krizek, Eric J. Miller, Srinivas Peeta, Thomas F. Rossi

Traveler Behavior and Values Committee
Eric J. Miller, University of Toronto (Chair), Thomas J. Adler, Rebekah Straub Anderson, Kay W. Axhausen, David Berrigan, Todd A. Brauer, Shaunna Kay Burbidge, Xinyu Cao, Cynthia Chen, Caspar G. Chorus, Kelly J. Clifton, Sean T. Doherty, Frederick W. Ducca, Dick Ferdinand Ettema, Joel E. Freedman, John P. Gliebe, Thomas F. Golob, Konstadinos G. Goulias, David T. Hartgen, Stephane Hess, Tae-Gyu Kim, Kevin J. Krizek, Vladimir Livshits, Hani S. Mahmassani, Abolfazl Mohammadian, Patricia L. Mokhtarian, Kai Nagel, Maren L. Outwater, Emily Parkany, Kimon E. Proussaloglou, David B. Reinke, Matthew J. Roorda, Darren M. Scott, Yoram Shiftan, Sivaramakrishnan Srinivasan, Peter R. Stopher, Ye Xin

Telecommunications and Travel Behavior Committee
Kevin J. Krizek, University of Colorado, Denver (Chair), Anne Aguilera, Eran Ben-Elia, Peter Bonsall, Colby M. Brown, Xinyu Cao, Juan Antonio Carrasco, Cynthia Chen, Sangho Choo, Martin Dijst, Frank P. Douma, Christopher E. Ferrell, Tae-Gyu Kim, Mei-Po Kwan, Taihyeong Lee, Glenn D. Lyons, Hani S. Mahmassani, Abolfazl Mohammadian, Patricia L. Mokhtarian, Ahmed Ibrahem Ibrahem Mosa, Sivaramakrishnan Srinivasan, Wei Tang, Edward Weiner

Peer review is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 2009.

Transportation Research Board Staff
Kimberly M. Fisher, Associate Division Director, Transportation Planning
Freda R. Morgan, Senior Program Associate
Mary Kissi, Senior Program Associate

Publications Office
Susan Fleshman, Editor; Elizabeth C. Horowitz, Production Editor; Deborah Uffelman, Proofreader; Scott E. Hitchcock, Manuscript Preparer

Ann E. Petty, Managing Editor; Juanita Green, Production Manager; Phyllis Barber, Publishing Administrator; Jennifer J. Weeks, Manuscript Preparation Manager

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

TRB SPONSORS*

Transportation Departments of the 50 States, Puerto Rico, and the District of Columbia

Federal Government

U.S. Department of Transportation
- Federal Aviation Administration
- Federal Highway Administration
- Federal Motor Carrier Safety Administration
- Federal Railroad Administration
- Federal Transit Administration
- National Highway Traffic Safety Administration
- Research and Innovative Technology Administration

Bureau of Indian Affairs
Science and Technology Directorate, U.S. Department of Homeland Security
U.S. Army Corps of Engineers
U.S. Coast Guard

Nongovernmental Organizations

American Association of State Highway and Transportation Officials
American Public Transportation Association
American Transportation Research Institute
Association of American Railroads

---

*As of August 2010.

TRANSPORTATION RESEARCH RECORD

Journal of the Transportation Research Board, No. 2157

Contents

Foreword

The 2010 series of the *Transportation Research Record: Journal of the Transportation Research Board* consists of approximately 900 papers selected from 3,700 submissions after rigorous peer review. The peer review for each paper published in this volume was coordinated by the committee acknowledged at the end of the text; members of the reviewing committees for the papers in this volume are listed on page ii.

Additional information about the *Transportation Research Record: Journal of the Transportation Research Board* series and the peer review process appears on the inside back cover. TRB appreciates the interest shown by authors in offering their papers, and the Board looks forward to future submissions.

Note: Many of the photographs, figures, and tables in this volume have been converted from color to grayscale for printing. The electronic files of the papers, posted on the web at **www.TRB.org/TRROnline**, retain the color versions of photographs, figures, and tables as originally submitted for publication.

Measurement Conversion Factors

To convert from the unit in the first column to the unit in the second column, multiply by the factor in the third column.

Customary Unit	*SI Unit*	*Factor*
Length		
inches	millimeters	25.4
inches	centimeters	2.54
feet	meters	0.305
yards	meters	0.914
miles	kilometers	1.61
Area		
square inches	square millimeters	645.1
square feet	square meters	0.093
square yards	square meters	0.836
acres	hectares	0.405
square miles	square kilometers	2.59
Volume		
gallons	liters	3.785
cubic feet	cubic meters	0.028
cubic yards	cubic meters	0.765
Mass		
ounces	grams	28.35
pounds	kilograms	0.454
short tons	megagrams	0.907
Illumination		
footcandles	lux	10.76
footlamberts	candelas per square meter	3.426
Force and Pressure or Stress		
poundforce	newtons	4.45
poundforce per square inch	kilopascals	6.89

Temperature

To convert Fahrenheit temperature (°F) to Celsius temperature (°C), use the following formula:
°C = (°F – 32)/1.8

SI Unit	*Customary Unit*	*Factor*
Length		
millimeters	inches	0.039
centimeters	inches	0.394
meters	feet	3.281
meters	yards	1.094
kilometers	miles	0.621
Area		
square millimeters	square inches	0.00155
square meters	square feet	10.764
square meters	square yards	1.196
hectares	acres	2.471
square kilometers	square miles	0.386
Volume		
liters	gallons	0.264
cubic meters	cubic feet	35.314
cubic meters	cubic yards	1.308
Mass		
grams	ounces	0.035
kilograms	pounds	2.205
megagrams	short tons	1.102
Illumination		
lux	footcandles	0.093
candelas per square meter	footlamberts	0.292
Force and Pressure or Stress		
newtons	poundforce	0.225
kilopascals	poundforce per square inch	0.145

Temperature

To convert Celsius temperature (°C) to Fahrenheit temperature (°F), use the following formula:
°F = (°C × 1.8) + 32

Abbreviations Used Without Definitions

AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACRP	Airport Cooperative Research Program
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASTM	American Society for Testing and Materials (known by abbreviation only)
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
ITE	Institute of Transportation Engineers
NASA	National Aeronautics and Space Administration
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
RITA	Research and Innovative Technology Administration
SAE	Society of Automotive Engineers
SHRP	Strategic Highway Research Program
TCRP	Transit Cooperative Research Program
TRB	Transportation Research Board

Spatial Econometrics Approach to Integration of Behavioral Biases in Travel Demand Analysis

Oleg A. Smirnov

Random utility models customarily assume strict independence of individual decision makers. Evidence of crowding, peer pressure, herd behavior, and other instances of spontaneous discrete choice coordination indicates that decision makers interact and thus affect choices made by others. Socially influenced individual choices become biased toward either agreeing with or contradicting the choices made by peers. Because many social interdependencies are spatial, a basic spatial discrete choice model was obtained by extending random utility theory to discrete choices made by heterogeneous spatially dependent individuals. Although interdependencies are unobserved, the model permits the study of behavioral biases arising from spatial interdependencies. The spatial discrete choice model is shown to address the effects of behavioral biases on conditional choice probabilities, the marginal effects of exogenous variables on revealed preferences, and the spatial patterns of discrete choices. A pseudo maximum likelihood (PML) estimator for the model is developed, and closed-form expressions for conditional choice probability estimates are derived. The PML estimator is shown to be consistent and computationally feasible for large spatial data sets. Simulated data were used to illustrate the performance of the PML estimator for the spatial discrete choice model.

The importance of spatial effects in travel demand is increasingly recognized in the literature (*1–8*). Spatial heterogeneity and spatial dependence are important types of spatial effects relevant to linear spatial autoregressive models (*9*). In spatial discrete choice models, there are three spatial effects: correlation, heterogeneity, and heteroskedasticity (*1*). Heteroskedastic effects play out in dramatic contrast to the way they do in spatial autoregressive models. In linear models, heteroskedasticity is an important econometric issue (*10*) but has no substantive effect on the dependent variable. In discrete choice models, heteroskedasticity has weighty significance; it affects individual choice probabilities, that is, it has no effect on the ordering of alternatives, but it does affect the probabilities with which specific alternatives are chosen. Furthermore, spatial heteroskedasticity is an important consequence of spatial dependence (*11*). For this reason, by accounting for unobserved spatial heteroskedasticity only and ignoring other artifacts of spatial dependence, one can arrive at the consistent likelihood-based estimator for the spatial discrete choice model.

Department of Economics, University of Toledo, 2801 West Bancroft Street, Toledo, OH 43606-3390. oleg.smirnov@utoledo.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 1–10.
DOI: 10.3141/2157-01

Spatial heterogeneity is an important spatial effect in the context of discrete choice models. It refers to the uneven spatial distribution of decision makers; that is, the differences between decision makers at different locations are fundamental, or unmeasurable. Spatial heterogeneity does not bring new concepts or issues concerning spatial behavior, because it is easily explained within the framework of independent decision makers. Both heterogeneous preferences and alternatives fit perfectly into the classical random utility model by including explanatory variables pertaining to the location of decision makers (*12–14*).

Spatial dependence is a particularly important spatial effect in spatial discrete choice models, because spatial interdependence of individuals presents a new aspect of individual behavior not found in models with noninteractive independent individuals. Spatial interdependencies between individuals affect their preferences, creating the phenomenon of socially influenced decision making, so that individuals neither behave strictly independently nor reach decisions jointly. Therefore, the solution to spatial dependence lies in spatial dependence in preferences, while the choices are still made by individual decision makers who are relying on the utility-maximizing principle. Social influence in decision making implies that individuals either agree or disagree with choices made by others more often than would be implied from independent decision making. Conceptually, generalized extreme value (GEV), GEV-based, and mixed logit models (*15–17*) are essentially single decision-maker models, because individuals are assumed to be independent, so that neither choices nor preferences of others affect the decisions of anyone else. This classical random utility model setting frequently appears to be implausible because individual decision makers routinely exchange opinions about their alternatives, share information about unobserved characteristics, and hypothesize about advantages or disadvantages of new routes, modes of transportation, and so forth. Individual decision making is a social process, in which interactions between individuals inadvertently become a part of the decision-making process. Assuming that casual socializing is mostly spatial because transaction cost and the relevance of the choice set impose implicit limitations on how far effective interactions can go, one can say that most social interactions are essentially spatial.

In the travel demand literature, spatial dependence has two distinctive aspects: spatial dependence in the choice set and spatial dependence in the set of preferences. The first aspect of spatial dependence is spatial correlation in the choice set, which is modeled by imposing the error structure within each decision maker's choice set. Forfeiting the ease of the analytical solution encountered in the multinomial logit model, spatial dependence among alternatives comfortably fits into the GEV framework, since individuals are

assumed to be independent. Reports of substantial progress in this area are available elsewhere (*2, 3, 6, 7*).

The second aspect of spatial dependence in discrete choice modeling is spatial dependence (both heteroskedasticity and correlation) in the set of preferences. This phenomenon is essentially behavioral, as it concerns the behavior of interacting decision makers. Whereas behavioral economics (*18*) studies specific mechanisms and various phenomena of behavioral biases (herd effect, peer pressure, information sharing, etc.), with space imposing an effective constraint on the scope of these effects, social influence is spatial and conveyed through neighbors. Behavioral bias toward agreeing with peers (neighbors) reflects a positive spatial dependence, whereas bias toward disagreeing signifies a negative spatial dependence between individual decision makers. There are important differences between spatial dependence and conditional choice probabilities. Although spatial dependence has no effect on individual choice probabilities, it does affect the joint probability distribution of choices made by individual decision makers. Ignoring spatial interdependencies in individual decision making leads to a biased assessment of individual preferences and inconsistent estimates of choice probabilities and model parameters. For example, sampled travel data can lead to inconsistent estimates because of potentially significant spatial interdependencies (*19*).

Spatial dependencies in models for transportation and residential choice location appear to deprive researchers of analytical, closed-form estimators of choice probabilities. Thus the recent literature focuses on simulation as a method for calibrating the models (*20–22*). Simulation requires an extensive team of experts for calibrating and analyzing the model, although spatial dependencies between decision makers are set aside. An analytically tractable estimator is important for model analysis and forecasting, since closed-form expression for the choice probability is easier to analyze and interpret. Difficulties in obtaining such an estimator pushed researchers to use spatial autoregressive models (*23*), in which the discrete nature of trip generation is effectively ignored. Although such a "something is better than nothing" solution allows one to detect spatial dependence, it cannot be transferred easily to discrete choice models, in which random utility is unobserved by a researcher. Another solution suggested in the literature is the use of alternative approaches to the notion of spatial dependence (*24, 25*), where the decision rule does not need to follow the random utility framework. Although such a shortcut might simplify computations, the interpretation of results and analysis of behavioral aspects become problematic because of the lost link between decision makers' choices and preferences.

The complicating issue in dealing with spatial dependence among decision makers from an econometric point of view is the presence of both spatial heteroskedasticity and spatial correlation between decision makers. The lack of substantial progress in this direction confirms that fitting spatial dependence into discrete choice models, or making a spatial extension of the GEV or mixed logit frameworks, is not an easy task (*25*). Incorporation of spatial dependence in the error term structure of the freight transportation model has resulted in a model with no analytically derived estimator (*26*). Approximation of spatial dependence in preferences with an easy-to-compute expression for the closed-form probability results in a model that is inconsistent with random utility theory (*25*).

The paper introduces the basic spatial discrete choice model and outlines the concept of the spatial multiplier matrix, an important aspect of the analysis of spatial dependence. Then an approach is developed to the analysis of behavioral biases in travel demand using estimates obtained from the pseudo maximum likelihood (PML) estimator. It shows that the aggregate and individual-level spatial effects can be quantified by using parameter estimates without requiring simulations. Simulated data are used to illustrate the method performance for various data set sizes and values of the spatial autoregressive coefficient. The results from spatial and non-spatial settings are contrasted, with empirical application as detailed by Smirnov and Egan (*27*).

SPATIAL ECONOMETRICS APPROACH TO DISCRETE CHOICE

A simple example can be used to illustrate the notion of spatial dependence in the discrete choice model. Suppose there are two neighbors, David and Frank. Each faces a choice set that contains two alternatives. Each alternative indicates an available option of urban commute. The first alternative is to drive a car, and the second alternative is to take the bus. Suppose David's choice probability of driving a car is $P_D(\text{car}) = .6$ and Frank's choice probability of driving a car is $P_F(\text{car}) = .5$. Respectively, David's probability to take the bus is $P_D(\text{bus}) = .4$, and Frank's probability to take the bus is $P_F(\text{bus}) = .5$. David and Frank drive their own cars and do not carpool. The concern here is the probability of congestion, which is the outcome when both David and Frank decide to drive. Relevant are not only the factors and conditions affecting David's and Frank's choices but also their joint probability distribution.

If David and Frank are independent in their decision making, the probability that both will drive a car, according to statistical theory, is equal to the product of their respective choice probabilities, so $\Pr(D = \text{car}, F = \text{car}) = P_D(\text{car}) * P_F(\text{car}) = .6 * .5 = .3$. Probabilities of all other combinations of David's and Frank's choices are shown in the first row of Table 1.

Now suppose David and Frank are not independent. The second row of Table 1 demonstrates joint choice probabilities for the case of positive spatial dependence. Although their individual choice probabilities are the same as in the first row, their joint probability distribution is different. Specifically, the probabilities that David and Frank will come up with identical decisions are higher than in the case of independent decision making. For instance, the probability that both will drive a car is .45 in the case of a positive spatial dependence, versus .30 when they are independent; the probability that both will take the bus is .35 versus .20. Thus, positive spatial dependence in this example does not in any way distort individual

TABLE 1 Joint Choice Probabilities

	Frank's Choice	
David's Choice	Drive a Car	Take the Bus
Independent Decision Making		
Drive a car	.30	.30
Take the bus	.20	.20
Positive Spatial Dependence		
Drive a car	.45	.15
Take the bus	.05	.35
Negative Spatial Dependence		
Drive a car	.20	.40
Take the bus	.30	.10

choice probabilities but does affect the probabilities of decision makers to agree on the choice.

The third row of Table 1 depicts the situation of negative spatial dependence. Again, the individual choice probabilities are unaffected by spatial dependence. Instead, the probabilities that decision makers will make identical choices are smaller than in the case of independent decision making. The probabilities of making different choices will be higher than in the case of spatial independence.

In the extreme case of positive spatial dependence, individual decision makers will converge on making identical choices, that is, both will choose to drive or both will choose to ride the bus. This situation implies that individual choice probabilities tend to be closer to each other. In the case of extreme negative spatial dependence, the probabilities of identical choices will be zero; that is, it is more likely that one will drive a car and the other ride the bus.

Extending the example to more than two individual decision makers requires definition of the structure of spatial dependencies between individuals and the way these dependencies affect individual choice probabilities.

Spatial Neighborhood

The structure of space is defined by the set of its elements and relations. In travel demand analysis, the basic element of space is a household, the number of which is finite. Each household is endowed with a unique location, which does not change during the analysis. The relations between households are fully defined by partial ordering, which is given by the set-theoretic notion of neighborhood.

Formally, let N be a finite set of households. The spatial neighborhood W is a subset of the product set $N \times N$. From a topological perspective, a spatial neighborhood is a binary relation, so that each pair of households $(i, j) \in N \times N$ is associated with either 1 (neighbors) or 0 (not neighbors): $W(i, j) = \{0|1\}$. Formally, spatial relation W is a mapping $W : N \times N \rightarrow \{0|1\}^{N \times N}$. This relation is symmetric $W(i, j) = W(j, i)$ and nonreflexive $W(i, i) = 0$ for $i \in N$. Households i and j are neighbors if and only if $W(i, j) = 1$.

A spatial neighborhood defines the range of spatial interdependencies, so that households directly interact only with households in their respective neighborhood. Each household is identified by its geographical location, and its spatial neighborhood is the set of households within close geographical proximity. The specific definition of neighborhood may vary by application, because it is important to select the definition of neighborhood that corresponds to the nature of interdependencies.

For a finite set of locations, the notion of spatial neighborhood is conveniently given by a spatial weights matrix W such that $W(i, j) = 1$ if i and j are neighbors and zero otherwise. Construction of a spatial weights matrix, however, requires the specification of a feasible neighborhood test or statistic, so that the pairs of neighbors can be arranged as a spatial weights matrix.

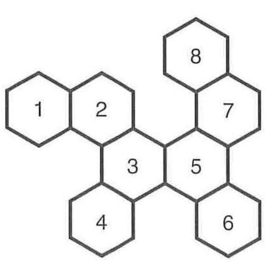


FIGURE 1 Example of spatial layout for which "neighborhood" is defined by common boundary between neighbors.

Specification of Spatial Weights by Using Residential Location Principle

For the spatial arrangement depicted in Figure 1, the spatial unit of observed data is given by polygons. Since there are eight polygons, the resulting spatial weights matrix is an 8×8 nonnegative matrix. Defining neighboring polygons as areas that share at least one point of common boundary leads to the following structure of spatial weights:

$$W = \begin{pmatrix} 0 & w_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ w_{21} & 0 & w_{23} & 0 & 0 & 0 & 0 & 0 \\ 0 & w_{32} & 0 & w_{34} & w_{35} & 0 & 0 & 0 \\ 0 & 0 & w_{43} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{53} & 0 & 0 & w_{56} & w_{57} & 0 \\ 0 & 0 & 0 & 0 & w_{65} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{75} & 0 & 0 & w_{78} \\ 0 & 0 & 0 & 0 & 0 & 0 & w_{87} & 0 \end{pmatrix}$$

By construction, nonzero spatial weights are assigned value 1, but to simplify model analysis and ensure spatial stationarity, spatial weights matrices are routinely row standardized. Row standardization involves scaling each row so that the sum of weights in the row is equal to 1. This operation has no effect on the topology of the spatial interdependencies or the structure of nonzero elements in the spatial weights matrix.

For example, in Figure 1 the first location has only one neighbor, Location 2; Location 2 has two neighbors, Locations 1 and 3; and so on. Therefore, there are only two nonzero entries in the second row of the spatial weights matrix: w_{21} and w_{23}. After row standardization, the corresponding elements in the first row of the matrix will have equal weights $w_{21} = w_{23} = 0.5$. Because of symmetry of spatial weights, there are exactly two nonzero elements in the second column of the spatial weights matrix: w_{12} and w_{32}. Row standardization ensures that the aggregate influence of each household's neighbors adds up to 1 regardless of the number of neighbors.

Specification of Spatial Weights by Using Transportation Network Principle

The interdependencies between individual decision makers are determined by the topology of the transportation network. For the network depicted in Figure 2, the usual graph-theoretic conventions

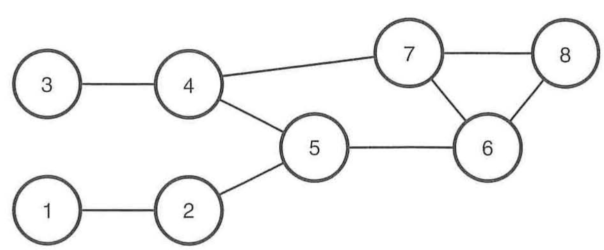


FIGURE 2 Example of spatial layout for which "neighborhood" is defined by transportation network.

apply. Each node corresponds to a decision maker, and each edge corresponds to a connection between a pair of individuals. Interactions occur only between directly connected individuals. For example, Individual 4 is connected to Individuals 3, 5, and 7. All other pairs of individuals are presumed to be noninteracting with each other. The corresponding structure of the spatial weights matrix for the example in Figure 2 is as follows:

$$W = \begin{pmatrix} 0 & w_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ w_{21} & 0 & 0 & 0 & w_{25} & 0 & 0 & 0 \\ 0 & 0 & 0 & w_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & w_{43} & 0 & w_{45} & 0 & w_{47} & 0 \\ 0 & w_{52} & 0 & w_{54} & 0 & w_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & w_{65} & 0 & w_{67} & w_{68} \\ 0 & 0 & 0 & w_{74} & 0 & w_{76} & 0 & w_{78} \\ 0 & 0 & 0 & 0 & 0 & w_{86} & w_{87} & 0 \end{pmatrix}$$

Initially, all connected pairs of individuals are assigned unadjusted spatial weights values of one. Row standardization is the procedure of scaling spatial weights by a positive value so that the sum of spatial weights in the row equals 1. For instance, after row standardization, the row corresponding to Individual 4 will have spatial weights $w_{43} = w_{45} = w_{47} = 1/3$. In practice, spatial weights matrix W might become quite large because its dimension is $n \times n$, where n is the number of decision makers.

Spatial Discrete Choice Model

Suppose the entire society is given by the set of individuals $N = \{1, 2, \ldots, n\}$. Each individual chooses one and only one alternative from the set $M = \{1, 2, \ldots, m\}$. Note that $N \cap M = \varnothing$, that is, alternatives in the choice set are not associated with other individuals. Spatial dependence between individuals in this paper is understood as the spatial dependence between individual preferences. In all other respects, decision makers are independent and their choices have no effect on the utility of other individuals—there are no externalities from the decision-making process. Individual preferences are modeled by utility, which is assumed to be additive. Denote u_{qj} random utility of individual q from selecting alternative j. Denote $u_j = (u_{1j}, u_{2j}, \ldots, u_{nj})^T$ the $n \times 1$ vector of individual random utilities from alternative j, $v_j = (v_{1j}, v_{2j}, \ldots, v_{nj})^T$ the vector of private deterministic components of the individual utilities, and $\epsilon_j = (\epsilon_{1j}, \epsilon_{2j}, \ldots, \epsilon_{nj})^T$ the vector of private stochastic components of individual utilities from alternative j.

The basic spatial random utility model is

$$u_j = \rho W u_j + v_j(\beta) + \epsilon_j \qquad j \in M \tag{1}$$

where $\rho \in \Theta_\rho$ and $\beta \in \Theta_\beta$ are model parameters. In this model, only linear effects of interactions are captured. The interactions are defined by the $n \times n$ nonnegative spatial weights matrix W. A zero entry in the matrix W indicates a noninteracting pair of individuals, and a positive $W_{tr} > 0$ indicates that individual r affects utility of individual t, that is, the marginal effect of individual r on t is $\partial u_{tj}/\partial u_{rj} = W_{tr}$ for all $j \in M$. The structure of nonzero weights is symmetric: $W_{ij} \neq 0 \Leftrightarrow W_{ji} \neq 0$. For simplicity of further analysis, the spatial weights matrix W is row standardized so that its largest eigenvalue is equal to 1.

With the row-standardized spatial weights matrix, the necessary condition for Model 1 to represent stationary spatial process, the coefficient ρ is bounded by the parameter domain $\Theta_\rho = (1/\omega, 1)$, where ω is the lowest eigenvalue of the spatial weights matrix W. This condition ensures that the transformation matrix $I - \rho W$ is nonsingular and its inverse $(I - \rho W)^{-1}$ is finite and positive definite. The parameter domain Θ_ρ allows for positive, negative, and zero values of the autoregressive coefficient ρ. If $\rho > 0$, interdependencies between individuals can be characterized as cooperative, and individual preferences exhibit substantial internal similarities. This effect leads to individuals being more likely to make choices identical to those made by neighbors, which will appear to be as a behavioral bias favoring choices made in the neighborhood. Negative ρ indicates individual preferences are repulsive as if individuals aim to underscore their individuality within the neighborhood by not following dominant preferences in the neighborhood. In this case, the choices will appear to be biased away from the choices followed by neighbors. A zero value of ρ suggests lack of substantial interdependence and, therefore, neighborhood-influenced behavioral bias.

The decision rule is assumed to be consistent with the notion of rationality of individual decision makers. Denote by $y_j = (y_{1j}, y_{2j}, \ldots, y_{nj})^T$ the $n \times 1$ vector of discrete choices with regard to alternative $j \in M$. The relation between individual utilities and chosen alternatives is

$$y_{qj} = \begin{cases} 1 & \text{if } u_{qj} \geq u_{qi} \qquad i \in M \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

so that one and only one alternative is chosen by each individual: $\sum_j y_{qj} = 1, \forall q \in N$.

Finally, stochastic components ϵ are assumed to be independently identically distributed and drawn from the extreme value Type I (Gumbel) distribution with the joint probability density function:

$$f(\epsilon_{11}, \ldots, \epsilon_{nm}) = \prod_{q=1}^{n} \prod_{j=1}^{m} \exp\left(-e^{-\epsilon_{qj}}\right) e^{-\epsilon_{qj}} \tag{3}$$

The combination of the spatial random utility model (Equation 1), the decision rule (Equation 2), and the probability measure (Equation 3) makes up the basic spatial discrete choice model.

The reduced form of the spatial random utility model is

$$u_j = Z v_j(\beta) + Z \epsilon_j \qquad j \in M \tag{4}$$

where $Z = (I - \rho W)^{-1}$ is the spatial multiplier matrix.

The role of the multiplier matrix Z in the spatial discrete choice model is important. The model as given by Equations 1 through 3 specifies the spatial distribution of the immediate effect of the shock ϵ_{sj} on random utility u_{qj}. It is equal to 1 for the individual q and zero for others. In contrast, the full spatial effect of the shock on random utilities is determined by the multiplier matrix $Z = (z_{qs})$. Specifically, the element z_{qs} indicates the full effect of the shock in random utility u_{sj} on the random utility u_{qj}, $\partial u_{qj}/\partial \epsilon_{sj} = z_{qs}$, $j \in M$. The multipliers are identical across alternatives, but values of multipliers generally vary across locations. This interactivity is nontrivially parameterized by the coefficient ρ, varies across pairs of locations, but is independent of the utility levels or alternatives. The interpretation of these multipliers

is relevant in the analysis of spatial interdependencies as it offers an alternative to social interactions analysis in the "linear-in-means" model (*28*) and similar settings (*29*).

Some properties of these multipliers are relevant for the analysis of spatial interactions:

1. The positive individual shock in the preferences of one individual has a nonnegative effect on the preferences of other individuals: $z_{qs} \geq 0$ for any pair of individuals q and s, as long as $\rho \geq 0$.
2. The effect of shock on the preferences of the individual where the shock originally occurred is greater than 1: $z_{ss} > 1$ for a well-connected spatial weights matrix and nonzero spatial dependence; this implies that the net benefit of social interactions for any individual is positive: $z_{ss} - 1 > 0$.
3. The multiplier effect for any pair of locations is weakly monotonically increasing for positive ρ.
4. The multiplier effect of observed shock is identical to that of unobserved shock; that is, the multiplier for the observed utility is identical to that for the unobserved utility.

Property 4 easily follows from inspection of the reduced form in Equation 4. The validity of Properties 1 through 3 is easy to establish by examining the Taylor series of the multiplier matrix:

$$Z = \left(I - \rho W\right)^{-1} = \lim_{K \to \infty} \sum_{t=0}^{K} \rho^t W^t$$

The convergence of the series is guaranteed by the limits on the eigenvalues of nonnegative matrix W.

One can conveniently decompose the matrix of full spatial effects from utility shock into private effects and social effects. Denote by D the diagonal matrix comprising diagonal elements of the multiplier matrix Z: $d_{ii} = z_{ii}$, $\forall\, i \in N$ and $d_{ij} = 0$, $i \neq j$. These elements represent full private effects of shocks in individual utility on the utility of that individual. Elements of matrix $Z - D$ represent social effects, which is the effect of shock in the individual utility on utilities of other individuals. Notice that both matrices are positive in the presence of spatial dependence.

LIKELIHOOD-BASED ESTIMATION

PML Estimator

Identification of the spatial discrete choice model for maximum likelihood estimation varies largely with the structure of observed utility $v_j(\beta)$. The common aspect of identification, however, is the conditions associated with spatial interdependencies between decision makers, including these three conditions: (*a*) bounded spatial autoregressive coefficient ρ so that the multiplier matrix Z is positive definite; (*b*) limitations on the topology of the nonzero elements of the spatial weights matrix–constant topological dimension (*30*) to satisfy asymptotic identification properties for spatial autoregressive models (*31*); and (*c*) the absence of disconnected individuals. In this analysis, the model is assumed to be identified and asymptotically identified.

The log likelihood function for the discrete choice model of Equations 1 through 3 is

$$L(\theta; y) = \ln P\left(y \middle| \beta, \rho\right) \tag{5}$$

where $P(y|\beta, \rho) = \Pr(Y_{1k(1)} = 1, Y_{2k(2)} = 1, \ldots, Y_{nk(n)} = 1|\beta, \rho)$ is the joint probability for the discrete random variable Y taken at y and $k(q)$ is the alternative chosen by the individual q. Denote indicator function $I(a): R \rightarrow \{0,1\}$, which is defined for any real argument and takes value 1 if a is positive and zero otherwise. The joint probability for Equation 5 is given by

$$P\left(y \middle| \beta, \rho\right) = \int \ldots \int \left(\prod_{q=1}^{n} \prod_{i=1}^{m} I\left(u_{qk(q)} \geq u_{qi}\right) \right) f_\epsilon\left(\epsilon_{11}, \ldots, \epsilon_{nm}\right) d\epsilon \tag{6}$$

where $d\epsilon = d\epsilon_{11} \ldots d\epsilon_{1m} \ldots d\epsilon_{n1} \ldots d\epsilon_{nm}$. By using joint probability to obtain the likelihood function and setting up the likelihood maximization problem, one obtains the maximum likelihood estimator. However, the term $(I - \rho W)^{-1}\,\epsilon$ is analytically intractable (*32*), which precludes the analytical formulation of the likelihood in Equation 5. The conventional approach involving calculation of $P(y|\beta, \rho)$ for any $(\beta, \rho) \in \Theta$ is challenging because the log likelihood in Equation 5 cannot be simplified into a summation of easier-to-compute integrals of smaller dimensions as typically is accomplished in practice in models without spatial interactions between individuals (*6, 17*). Sheer dimension of the computational problem makes it impractical if not impossible to evaluate Equation 6 through use of numerical integration.

The starting point for the PML estimator is the reduced form of Model 4. Denote $n \times n$ matrix D that consists of diagonal elements of the matrix Z. Then diagonal matrix D indicates private effects of random shocks on the individual utilities, and the matrix $Z - D$ is the matrix of social components of spatial multiplier, that is, the effects of the shocks in the individual utility on the utilities of other individuals in the society.

The conditional choice probability for the individual q to select an alternative j is:

$$P_{qj} = \Pr\left(u_{qj} \geq u_{qi}, i \in M\right) = P\left(y_{qj} = 1 \middle| \left\{\epsilon_{si}, s \in \left(N \setminus q\right), i \in M\right\}\right) \tag{7}$$

From Model 4, the random utility is

$$u_j = Zv_j\left(\beta\right) + Z\epsilon_j = Zv_j\left(\beta\right) + \left(Z - D\right)\epsilon_j + D\epsilon_j \tag{8}$$

Since $d_{qq} \geq 1$, one can safely scale the vector of random utilities by the diagonal elements, that is, $\Pr(u_{qj} \geq u_{qi}) = \Pr(u_{qj}/d_{qq} \geq u_{qi}/d_{qq})$. Denote the following:

$$g_{qj} = \sum_{t=1}^{n} z_{qt} v_{tj}\left(\beta\right) + \sum_{t=1}^{n} \left(z_{qt} - d_{qt}\right)\epsilon_{tj}$$

where $t \in N$ indicates a column of the spatial multiplier matrix. Then the scaled random utility is

$$\frac{u_{qj}}{d_{qq}} = \frac{g_{qj}}{d_{qq}} + \epsilon_{qj} \tag{9}$$

Notice that random components in scaled random utility in Equation 9 are independently identically distributed across alternatives for each individual. The diagonal elements in the matrix $Z - D$ are 0. Therefore, random variables g_{qj} and ϵ_{qj} are statistically independent. This independence allows one to use GEV model formulation to

obtain individual conditional probabilities for scaled random utility (9) as follows:

$$P_{qj} = \frac{\exp\left(\dfrac{g_{qj}}{d_{qq}}\right)}{\sum_{i=1}^{m}\exp\left(\dfrac{g_{qi}}{d_{qq}}\right)} \tag{10}$$

In the full-information maximum likelihood estimation, conditional probabilities P_{qj} in Equation 10 are random. However, suppose each individual substitutes random social effects of the shocks in another individual's preferences with expectations of these shocks but still accounts for the full effects of the private shock. This assumption is justified as long as each individual q is capable of relating the effect of the random shock in his utility ϵ_{qj} to the enjoyed random utility u_{qj}, which is equal to the element (q, q) of the matrix Z: $\partial u_{qj}/\partial\epsilon_{qj} = d_{qq}$. This effect is identical across alternatives but varies across individuals. In a society with N individuals, random shocks in the utilities of all other individuals could be a burdening task, so substituting social effect with its expectation is the behavioral adaptation simplifying individual decision making. The expectation of aggregate social effect on individual q is

$$E_{\epsilon^-}\left[\frac{g_{qj}}{d_{qq}}\right] = \sum_{t=1}^{n} z_{qt}\frac{v_{tj}(\beta)}{d_{qq}}$$

where ϵ^- denotes the set of ϵ's from which qth component is removed. Substitution of this expression in Equation 10 yields the closed form for the expectation of conditional probabilities:

$$\hat{P}_{qj} = \frac{\exp\left(\sum_{t=1}^{n} z_{qt}\dfrac{v_{tj}(\beta)}{d_{qq}}\right)}{\sum_{i=1}^{m}\exp\left(\sum_{t=1}^{n} z_{qt}\dfrac{v_{ti}(\beta)}{d_{qq}}\right)} \tag{11}$$

The important distinction from the full maximum likelihood is that expectations of conditional probabilities $\hat{P}$ are nonrandom and the closed form in Equation 11 is nonrecursive. This greatly simplifies the log likelihood function based on the choice probability in Equation 6, which is

$$\hat{L}(\beta,\rho) = \sum_{q=1}^{n}\sum_{j=1}^{m} y_{qj}\ln\hat{P}_{qj} \tag{12}$$

and the PML estimator of β and ρ is the extremum estimator $\hat{\theta}_{\text{PML}} = (\hat{\beta}_{\text{PML}}, \hat{\rho}_{\text{PML}})$ that maximizes the log likelihood function in Equation 12. Since some relevant information about dependencies between probabilities is ignored, PML estimator $\hat{\theta}_{\text{PML}}$ may not be asymptotically efficient. Details of the estimation method are available elsewhere (8).

The PML estimator introduced here is equivalent to the maximum likelihood estimator for the spatial discrete choice model with spatial random utility:

$$\tilde{u}_j = Z(\rho)v_j(\beta) + D(\rho)\epsilon_j \tag{13}$$

The auxiliary model comprising Equations 13, 2, and 3 has the same observed deterministic components of individual random utilities as the original model of Equations 1 through 3. Private components of the unobserved spatial interdependencies $D(\rho)\epsilon_j$ in random utilities are identical in both models. However, error terms in the auxiliary model are independent, which substantially simplifies its maximum likelihood estimation. Since some of the information about effects of individual interdependencies is not present in the auxiliary model, parameter estimates obtained from PML would not be as efficient as those of the true maximum likelihood estimator if one were available. Computational details involving matrix transformations are available elsewhere (8).

Biases in Travel Demand Analysis Arising from Spatial Dependence

The first important distinction of the spatial discrete choice model from the incorrectly specified nonspatial model is that the log odds

$$\ln\frac{\hat{P}_{qj}}{\hat{P}_{qk}} = \ln\frac{\exp\left(\sum_{t=1}^{n} z_{qt}\dfrac{v_{tj}(\beta)}{d_{qq}}\right)}{\exp\left(\sum_{t=1}^{n} z_{qt}\dfrac{v_{tk}(\beta)}{d_{qq}}\right)} = \frac{1}{d_{qq}}\sum_{t=1}^{n} z_{qt}\left(v_{tj}(\beta) - v_{tk}(\beta)\right) \tag{14}$$

are given as a function of the spatial multiplier elements d_{qq}. Generally, elements d_{qq} vary across q; thus the model implies distinctive log odds across individuals even if they are exposed to identical spatial arrangements. In contrast, in the nonspatial model, coefficients d_{qq} are not present because there is no spatial heteroskedasticity in such a model. This shows that the ignoring of spatial interdependencies results in biased conditional probabilities because generally $d_{qq} \neq 1$. In addition, the average of d_{qq} is strictly greater than 1 in the presence of spatial dependence, thus yielding a systematic bias if ignored.

The analysis also highlights that spatial heteroskedasticity originates in spatial dependence but has behavioral connotations. Nonspatial heteroskedasticity—heteroskedasticity in part unrelated to spatial dependence—is easily accommodated within the GEV framework by appropriate scaling of random utility. However, it is justified in practice only if there is evidence of the presence of various levels of uncertainty for different decision makers. For example, a decision maker in a hurricane-affected area is, ceteris paribus, more affected by changes in weather than a decision maker who enjoys a more secure location. Uncertainty associated with weather thus will affect their respective utilities regardless of whether these individuals interact, which fits routinely into GEV framework. Heteroskedasticity in part related to spatial dependence emerges as the result of increased or decreased uncertainty because of spatial interdependencies. In a spatial setting, heteroskedasticity occurs naturally because better-informed or better-connected individuals will face the level of uncertainty that differs from that of less-informed decision makers. It is intuitive that more-connected individuals are more likely to follow the crowd, or agree with choices made by others, in either socially desirable or implausible directions. Poorly connected individuals are more likely to remain independent in their choices. For this reason, better-connected individuals are less likely to make choices predicted by the observed utilities. Other things being equal, networks with many pairs of interdependent individuals are less predictable than those with fewer connections because crowd impulses such as panic and apathy will have more powerful effects than when individuals are independent.

The second key aspect of the spatial discrete choice model is relevant for policy analysis and evaluation of the impact of the exogenous variables on individual choices. Specifically, in the nonspatial model the marginal effects of the change in exogenous variables

on individual utilities and, therefore, choice probabilities are defined by model parameters β. In contrast, in the spatial discrete choice model, the marginal effect of exogenous variables is spatially distributed. Hence, some individuals can be affected more than others. In the spatial discrete choice model, the multiplier matrix $Z(\rho)$ determines the spatial distribution of marginal effects. It can be shown that the marginal effects are identical to those in the nonspatial model only in exceptional circumstances. Generally,

$$E\left[\frac{\partial u_j}{\partial \beta}\right] = Z\frac{\partial v_j(\beta)}{\partial \beta} \tag{15}$$

In Equation 15, the spatial multiplier matrix $Z = (I - \rho W)^{-1}$ is equal to the identity matrix only when $\rho = 0$, which corresponds to the case of spatial independence; hence, the classical discrete choice model. In all other cases, the multiplier matrix Z determines the value of the effects. It is clear that in the presence of positive spatial dependence, the multiplier matrix is componentwise greater than the identity matrix.

The third pivotal characteristic of the spatial discrete choice model is its potential for analyzing dependencies and regularities in travel demand arising from characteristics of transportation networks. Use of the spatial discrete choice model to frame behavioral biases in travel demand is important to analysis of spatial patterns of network use and traffic congestion, as well as for forecasting demand for transportation infrastructure. That is, the multiplier matrix Z is susceptible to change in the value of the parameter ρ as in Equation 16:

$$\frac{\partial Z}{\partial \rho} = (I - \rho W)^{-1} W (I - \rho W)^{-1} \tag{16}$$

and the spatial weights matrix W as in Equation 17:

$$\frac{\partial Z}{\partial W} = \rho W \otimes (I - \rho W)^{-1} \tag{17}$$

Whereas Equation 16 is important for studying the effect of spatial dependence on choice probabilities, Equation 17 indicates the effect of changes in the structure of spatial dependencies. If those dependencies are driven primarily by the topology of transportation networks, Equation 17 indicates changes in the spatial multiplier matrix associated with improvements in the transportation network. The key aspect of this analysis, and its principal prerequisite, is the obtaining of parameter estimates $\hat{\rho}$ and $\hat{\beta}$. Computation of terms containing the multiplier matrix $Z = (I - \rho W)^{-1}$ for large data sets without inverting the matrix can be done efficiently through use of the sparse conjugate gradient method introduced in an earlier paper (*33*).

METHOD PERFORMANCE: ILLUSTRATION WITH SIMULATED DATA

For an illustration of finite sample properties of the estimation method, consider the spatial discrete choice model with random utility.

$$u_j = \rho W u_j + x_j\beta_1 + x_{qj}\beta_2 + \epsilon_j \tag{18}$$

where

j = travel alternative,
x_{qj} = travel cost for the individual q using alternative j, and
x_j = attribute of alternative j, which is uniform for all individuals.

Individuals are assigned to locations; the neighborhood structure is given by the spatial weights matrix W. To generate data for a simulations study, the spatial arrangement of individuals is assumed to follow the hexagonal pattern characteristic of the central place theory, where locations are given by regular hexagons similar to those depicted in Figure 1. Hexagons with common boundaries are assumed to be neighbors. To randomize the simulated grid, a fixed fraction of neighborhood ties (about 30% of all pairs) are severed. However, it is ensured that each location has at least one neighbor. As a result, the grid is irregular and random and encounters fewer nonzero elements in the spatial weights matrix than the regular grid would.

The simulated values are obtained as follows. Variables x are sampled from the uniform random distribution, so that simulated values are independently identically distributed. This is sufficient to satisfy regularity assumptions ensuring consistency of the PML estimator. Values of model parameters are set as $\beta_1 = 1$ and $\beta_2 = -2$ in all trials. The error terms are sampled from the Type I extreme value distribution. Then the conjugate gradient method is used to compute $u = (I - \rho W)^{-1}(x_1\beta_1 + x_2\beta_2 + \epsilon)$ for each alternative. Values of utilities are used to determine the alternative with the largest utility. For such an alternative for every individual q, the variable y_{qj} is set to 1; all other y_{qi}, $i \neq j$ are set to zero. The variables y, x_1, x_2 and the spatial weights matrix W are used as the input for the PML estimator in all trials.

The coefficient ρ varies, as shown in the tables. In all the trials there are eight alternatives. As shown in Table 2, in the correctly specified model the estimates of the spatial autoregressive coefficient are most accurate—the bias and mean squared error (MSE) are smallest across all coefficients. The bias of all three estimates shows a tendency to decline with the number of trials. The MSEs for all three estimates appear to be indifferent to the number of trials. The combination of these two properties is a necessary characteristic of a consistent estimator. The restricted model is obtained by setting $\rho = 0$ during estimation. This is equivalent to assuming that the discrete choice model is nonspatial; however, there is significant positive spatial dependence between decision makers. The estimates of coefficients β_1 and β_2 are biased and show no signs of convergence to true values for either of the coefficients.

There are two tendencies in the biases of coefficient estimates. First, bias (in absolute value) increases with the strength of spatial dependence. A closer look at the biases and MSEs of estimates of β_1 and β_2 indicates the differences in magnitudes of biases between them. Specifically, the second distinctive pattern in biases is that for a small spatial dependence $\rho = 0.05$ the bias in β_1 is greater than that in β_2. However, for stronger spatial dependence, bias in estimate of β_2 becomes larger in absolute value than that of β_1. The same pattern occurs for MSEs. This phenomenon indicates that the spatial dependence not accounted for, since it is more strongly correlated with travel cost, has a stronger effect on the travel cost coefficient β_2. Consequently, the estimate of β_1 is more biased and has larger variance. That is, the foremost consequence of ignoring spatial dependence in discrete choice models is biased and noisy estimates of travel cost. A secondary consequence, smaller but still significant, is the bias in substantive alternative-specific coefficients.

Another aspect of the PML estimator is its properties for large data sets. As established earlier, the PML estimator need not be efficient. The significance of this inefficiency should not be overvalued. As results in Table 3 indicate, the bias and MSE of the parameter estimates have a strong tendency to decline with the sample size. The MSE of the spatial autoregressive coefficient declines much faster with the sample size than does the MSE of β. This suggests that for large

TABLE 2 Accuracy of Parameter Estimates

	ρ		β_1		β_2	
Trial	Bias	MSE	Bias	MSE	Bias	MSE
Unrestricted Model: Spatial Dependence Estimated from Data						
$\rho = 0.05$						
5	0.0207	0.00082	−0.080	0.021	0.105	0.014
25	0.0053	0.00081	0.027	0.011	−0.015	0.024
100	−0.0032	0.00109	−0.013	0.013	0.037	0.026
$\rho = 0.20$						
5	0.0160	0.00100	−0.035	0.102	0.064	0.034
25	0.0005	0.00031	−0.022	0.011	0.025	0.037
100	−0.0018	0.00063	−0.004	0.0090	0.0040	0.034
$\rho = 0.50$						
5	0.0103	0.00049	0.0412	0.0042	−0.083	0.0166
25	−0.00093	0.00041	0.0379	0.0105	−0.069	0.0403
100	−0.00058	0.00044	0.011	0.018	−0.036	0.061
Restricted (Nonspatial) Model						
$\rho = 0.05$						
5	−0.05	N/A	−0.100	0.015	0.067	0.016
25	−0.05	N/A	−0.045	0.008	−0.022	0.024
100	−0.05	N/A	−0.054	0.013	0.007	0.028
$\rho = 0.20$						
5	−0.20	N/A	−0.089	0.013	−0.372	0.146
25	−0.20	N/A	−0.138	0.027	−0.313	0.108
100	−0.20	N/A	−0.118	0.026	−0.332	0.128
$\rho = 0.50$						
5	−0.50	N/A	0.083	0.0105	−1.083	1.173
25	−0.50	N/A	0.095	0.0124	−1.07	1.147
100	−0.50	N/A	0.110	0.014	−1.087	1.182

NOTE: Spatial data set comprised of 400 locations. Bias is the average bias of the estimate over the given number of trials. MSE is the mean squared error of the estimator.

TABLE 3 Accuracy of Method

	ρ		β_1		β_2	
Trial	Bias	MSE	Bias	MSE	Bias	MSE
Size of Data Set Is $N = 1{,}000$						
5	0.00143	0.00041	0.0326	0.0100	−0.0474	0.0121
25	−0.00076	0.00016	−0.0154	0.0072	0.0171	0.0151
100	−0.00269	0.00025	−0.00205	0.0032	0.0064	0.0089
$N = 5{,}000$						
5	0.00139	$4*10^{-5}$	0.0199	0.0013	−0.0386	0.0033
25	−0.00052	$5*10^{-5}$	−0.00091	0.00079	−0.00456	0.00247
100	−0.00127	$6*10^{-5}$	0.0029	0.0007	−0.00778	0.00279
$N = 20{,}000$						
5	−0.00015	$1*10^{-5}$	0.0019	0.00014	−0.0097	0.00075
25	0.0013	$1.6*10^{-5}$	0.0087	0.00033	−0.011	0.00086
100	−0.00038	$1.3*10^{-5}$	0.00062	0.00021	−0.014	0.00076

NOTE: In all trials, spatial autoregressive coefficient is $\rho = 0.20$.

spatial data sets, common in empirical analysis, efficiency of the PML estimate of the spatial autoregressive coefficient is not important. However, the bias and MSE of both estimates β_1 and β_2, although tending to decline with sample size, are higher, given that all coefficients in the simulations are of the same order of magnitude. Consequently, the bias and noisiness of estimates resulting from ignoring spatial dependence in discrete choice models is much more significant than the potential loss of efficiency in the PML estimator associated with a simplified log likelihood expression. An example of the application of this approach to recreational travel demand is available elsewhere (*27*).

CONCLUSION

The unifying feature of discrete choice models is their focus on the behavioral aspect of individual decision making. In the simplest form of these models, individual decision makers are independent and noninteractive, so individual choices are unrelated to choices of other decision makers in the society. Development of applied discrete choice models for spatial data requires a framework that allows for spatial effects—spatial dependence and spatial heterogeneity. Spatial dependence has two facets: dependence in the choice set and dependence in the set of preferences. Whereas spatial heterogeneity and spatial dependence in the choice set fit in the classical random utility framework, spatial dependence in the set of preferences is the key aspect of the spatial discrete choice model. Conceptually, dependencies between the preferences of individual decision makers are intrinsically behavioral, so spatial discrete choice is advocated as an approach to studying behavioral biases that result from spatial dependence.

Spatial dependence in the set of preferences is the spatial interdependence between the preferences of individual decision makers. This phenomenon leads to socially influenced decision making, so that individuals neither are fully independent nor reach decisions jointly. However their preferences are affected, each individual in the model follows the utility-maximizing principle, which allows substantive study of the behavioral biases that result from spatial dependence.

From an econometric point of view, spatial dependence in the set of preferences encounters spatial heteroskedasticity and spatial correlation. One can ignore the correlation to obtain the PML estimator that is based on the closed-form expression for conditional choice probabilities. This estimator is shown to be consistent but not necessarily asymptotically efficient. The use of closed-form probabilities is important for the analysis of behavioral biases. This result is important for model analysis and forecasting, since easy-to-use, closed-form probabilities for models of travel demand with spatial dependence have not been developed previously.

The magnitude of behavioral biases in the spatial discrete choice model is illustrated by use of the Monte Carlo study. It is shown that the classical random utility model is biased and inconsistent in the presence of spatial dependence. Transport costs, if present in the model, tend to be affected mostly in the misspecified model, since transport costs in this case are proxy for distance, which is an important delimiter of spatially dependent decision makers. Parameter estimates in the spatial model substantially differ from those of the nonspatial model, which suggests that behavioral biases in travel demand models are statistically significant even when spatial dependence is small (0.10) to moderate (0.30).

ACKNOWLEDGMENTS

The author thanks the referees for helpful comments. This research was sponsored in part by the National Center for Freight and Infrastructure Research and Education at the University of Wisconsin.

REFERENCES

1. Bhat, C., and H. Zhao. The Spatial Analysis of Activity Stop Generation. *Transportation Research Part B,* Vol. 36, No. 6, 2002, pp. 557–575.
2. Bhat, C. R., and J. Guo. A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. *Transportation Research Part B,* Vol. 38, No. 2, 2004, pp. 147–168.
3. Koppelman, F. S., and V. Sethi. Incorporating Variance and Covariance Heterogeneity in the Generalized Nested Logit Model: An Application to Modeling Long Distance Travel Choice Behavior. *Transportation Research Part B,* Vol. 39, No. 9, 2005, pp. 825–853.
4. Louviere, J., K. Train, M. Ben-Akiva, C. Bhat, D. Brownstone, T. A. Cameron, R. T. Carson, J. R. Deshzo, D. Fiebig, W. Green, D. Hensher, and D. Waldman. Recent Progress on Endogeneity in Choice Modeling. *Marketing Letters,* Vol. 16, No. 3–4, 2005, pp. 255–265.
5. Mohammadian, A., M. Haider, and P. S. Kanaroglou. Incorporating Spatial Dependencies in Random Parameter Discrete Choice Models. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005.
6. Sivakumar, A., and C. R. Bhat. Comprehensive, Unified Framework for Analyzing Spatial Location Choice. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 103–111.
7. Bekhor, S., and J. N. Prashker. GEV-Based Destination Choice Models that Account for Unobserved Similarities Among Alternatives. *Transportation Research Part B,* Vol. 42, No. 3, 2008, pp. 243–262.
8. Smirnov, O. A. Modeling Spatial Discrete Choice. *Regional Science and Urban Economics* (in press).
9. Anselin, L. *Spatial Econometrics: Methods and Models,* Kluwer, Dordrecht, Netherlands, 1988.
10. Kelejian, H. H., and D. P. Robinson. A Suggested Test for Spatial Autocorrelation and/or Heteroskedasticity and Corresponding Monte Carlo Results. *Regional Science and Urban Economics,* Vol. 28, No. 4, 1998, pp. 389–417.
11. McMillen, D. P. Probit with Spatial Autocorrelation. *Journal of Regional Science,* Vol. 32, No. 3, 1992, pp. 335–348.
12. Dugundji, E. R., and J. L. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1921,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.
13. Walker, J. L., and J. Li. Latent Lifestyle Preferences and Household Location Decisions. *Journal of Geographical Systems,* Vol. 9, No. 1, 2007, pp. 77–101.
14. Dugundji, E. R., and L. Gulyás. Sociodynamic Discrete Choice on Networks in Space: Impacts of Agent Heterogeneity on Emergent Outcomes. *Environment and Planning B: Planning and Design,* Vol. 35, No. 6, 2008, pp. 1028–1054.
15. McFadden, D. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Residential Location* (A. Karlqvist et al., eds.), North–Holland, Amsterdam, Netherlands, 1978, pp. 75–96.
16. Ben-Akiva, M. E., and S. R. Lerman. Discrete Choice Analysis: Theory and Application to Travel Demand. *MIT Press Series in Transportation Studies,* Vol. 9, MIT Press, Cambridge, Mass., 1985.
17. Train, K. E. *Discrete Choice Methods with Simulations.* Cambridge University Press, New York, 2003.
18. Pesendorfer, W. Behavioral Economics Comes of Age: A Review Essay on *Advances in Behavioral Economics. Journal of Economic Literature,* Vol. 44, Sept. 2006, pp. 712–721.
19. Bierlaire, M., D. Bolduc, and D. McFadden. The Estimation of Generalized Extreme Value Models from Choice-Based Samples. *Transportation Research Part B,* Vol. 42, No. 4, 2008, pp. 381–394.
20. Abraham, J. E., and J. D. Hunt. Random Utility Location, Production, and Exchange Choice; Additive Logit Model; and Spatial Choice

Microsimulations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2003,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 1–6.

21. Eluru, N., A. R. Pinjari, J. Y. Guo, I. N. Sener, S. Srinivasan, R. B. Copperman, and C. R. Bhat. Population Updating System Structures and Models Embedded in the Comprehensive Econometric Microsimulator for Urban Systems. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2076,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 171–182.
22. Zhang, K., H. S. Mahmassani, and C.-C. Lu. Probit-Based Time-Dependent Stochastic User Equilibrium Traffic Assignment Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2085,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 86–94.
23. Kim, J.-H., J.-H. Chung, I. Kim, and U. Lim. Spatial Dependence in Trip Generation Models: Spatial Econometric Perspective. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
24. Klier, T., and D. P. McMillen. Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples. *Journal of Business and Economic Statistics,* Vol. 26, No. 4, pp. 460–471.
25. Bhat, C. R., and I. N. Sener. Copula-Based Closed-Form Binary Logit Choice Model for Accommodating Spatial Correlation Across Observational Units. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
26. Garrido, R. A., and M. Leva. Port of Destination and Carrier Selection for Fruit Exports: A Multi-dimensional Space-Time Multinomial Probit Model. *Transportation Research Part B,* Vol. 38, No. 7, 2004, pp. 657–667.
27. Smirnov, O. A., and K. J. Egan. *Spatial Random Utility Model with an Application to Recreation Demand.* Department of Economics, University of Toledo, Ohio, 2009.
28. Graham, B. S. Identifying Social Interactions Through Conditional Variance Restrictions. *Econometrica,* Vol. 76, No. 3, 2008, pp. 643–660.
29. Hartmann, W. R., P. Manchanda, H. Nair, M. Bothner, P. Dodds, D. Godes, K. Hosanagar, and C. Tucker. Modeling Social Interactions: Identification, Empirical Methods and Policy Implications. *Marketing Letters,* Vol. 19, No. 3–4, 2008, pp. 287–304.
30. Smirnov, O. A., and L. E. Anselin. An $O(N)$ Parallel Method of Computing the Log-Jacobian of the Variable Transformation for Models with Spatial Interaction on a Lattice. *Computational Statistics and Data Analysis,* Vol. 53, No. 8, 2009, pp. 2980–2988.
31. Lee, L.-F. Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica,* Vol. 72, No. 6, 2004, pp. 1899–1925.
32. Anselin, L. Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics,* Vol. 27, No. 3, 2002, pp. 247–267.
33. Smirnov, O. Computation of the Information Matrix for the Models of Spatial Interaction on a Lattice. *Journal of Computational and Graphical Statistics,* Vol. 14, No. 4, 2005, pp. 910–927.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Empirical Case Study of Spatial–Temporal Student Activity Population

Jin Ki Eom, John R. Stone, and Kyungwoo Kang

The spatial–temporal activity–presence approach is used to estimate and validate hourly activity populations for individual buildings. A case study modeled the spatial–temporal activity of students at North Carolina State University. For the validation of results, student registration records provide observations of class schedules and locations for dynamic, time-varying class or study activity populations at individual buildings. Results show that the spatial–temporal activity–presence approach provides reliable estimates. This empirical case study should improve acceptance of the spatial–temporal activity procedure for activity distribution and demonstrate its value for other planning applications.

Development of activity-based approaches for travel demand modeling has advanced rapidly in the last decade. The approach varies depending on model components, data structures, and the goals of model applications (*1, 2*). Some activity-based approaches, such as those for Boise, Idaho (1994), New Hampshire (1996), and Stockholm, Sweden (2000), may replace conventional travel demand models for model accuracy and ability to address important transport polices such as travel demand management (*3*). However, the complexity of the structure of activity-based travel demand models hinders implementation in practice. If the activity-based travel demand model is to be more widely used (especially by metropolitan planning organizations), the procedures and models must be more straightforward, demonstrated, and validated—goals for this paper and for future research (*4*).

This study provides empirical evidence of model accuracy by implementing an approach of spatial and temporal activity presence for a university campus. Specifically, the empirical validation of estimated dynamic activity population at the building level is modeled and validated.

A university campus is an appropriate place for a case study that estimates and validates dynamic activity population (e.g., the number of people participating in a certain activity at a certain time of day), because the geographic boundaries of buildings and the associated populations are reliably identified by student registration records. (Activity population and activity presence have the same meaning for the spatial–temporal activity-based approach, and they are used interchangeably in this paper.) Registration records document hourly student population for each university building for each class activity. However, a university campus and its associated class activities have a narrower scope than do an urban area and its activities, and the validation is made at the building level, which is smaller than an urban traffic analysis zone (TAZ). However, validation of the estimated number of students engaged in an activity at particular buildings throughout a day compared to the actual number of students based on easily available registration records is relatively robust.

J. K. Eom, Railway Transport and Logistics Research Department, Korea Railroad Research Institute, Woulam-Dong, Uiwang, Kyonggi, 437-757, South Korea. J. R. Stone, Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Campus Box 7908, Raleigh, NC 27695-7908. K. Kang, Department of Transportation Engineering, University of Hanyang, 1271 Sa-1-dong, Ansan, 425-791, South Korea. Corresponding author: J. K. Eom, jkom00@krri.re.kr.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 11–21.
DOI: 10.3141/2157-02

LITERATURE REVIEW

The spatial–temporal activity approach is based on an innovative idea of Hägerstrand, whose time–space prism describes when people are engaged in different activities, the interrelationships between activities, and the temporal and spatial constraints (*5*), and on the basic idea that an individual's trip-making behavior is derived from the desire to participate in activities (*6*).

There are several classes of activity-based approaches:

- Computational process approaches: PCATS by Kitamura et al. (*7*) and FAMOS by Pendyala and Bhat (*4*),
- Rule-based approaches: Albatross by Arentze and Timmermans (*8*),
- Random utility–based models: Portland model (*9*), and
- Activity–presence–based approaches: CentreSIM (*10–12*).

PCATS (prism-constrained activity–traveler simulator) combines Hägerstrand's concept of a time–space prism of activity engagement and travel simulation. In the model, activities are classified as fixed or flexible. The time–space prisms are developed on the basis of the speed of travel, locations, and timing of the fixed activities in which the flexible activities are constrained. Within time–space prisms, dynamic activity engagement is obtained in time segments and geographic boundaries (*7*).

Alam and Goulias developed an emergency evacuation management system for the University Park campus of Pennsylvania State University (*13*). This system incorporates the behavioral patterns of individuals and takes into account local land use patterns at a microscopic level to predict individual activity participation in space and time. This approach was expanded to a regional travel demand modeling system called CentreSIM (*10*), which implemented the spatial–temporal activity–presence–based modeling approach for Center County, Pennsylvania. This approach assigns an individual to a spatial location according to land use (e.g., business information) and assigns a series of activities at the same time (*11*). Hourly activity population is estimated at the TAZ level, which allows the

approach to be compatible with a conventional travel demand model. Travel demand is estimated by calculating the difference of activity population in each time frame.

Kwan introduced three-dimensional geovisualization methods based on geographic information systems (GIS) for spatial–temporal dimensions of activity-travel patterns (*14*). Travel diary data for Portland, Oregon, were reproduced by using three-dimensional GIS techniques developing the temporal activity dimension and its interaction within a frame of daily space–time trajectories. Buliung and Kanroglou developed a prototype object-oriented GIS system to support exploration of household-level activity behavior and accomplished a case study that generates space–time trajectories from household-level geographic and activity-timing properties obtained from the Portland activity survey (*15*).

Beckx et al. estimated a dynamic activity-based population to evaluate exposure to air pollution (*16*). Within the Albatross modeling framework, a scheduling process developed the dynamic activity population present in hourly cross sections of the population of the city center of Utrecht, the Netherlands, to examine the inflow and outflow of people. Beckx et al. developed a geographic representation of the number of people present by time of day in the city center on an average weekday to analyze dynamic exposure to air pollution.

These applications of activity-based modeling demonstrate spatial–temporal activity allocation and use GIS techniques to visually represent activity travel behavior. However, these approaches do not clearly demonstrate how well activity-based approaches replicate actual dynamic activity population by comparing model estimates with real observations of activity population.

This study of a university campus is the first valid comparison of estimates to observed activity populations. The paper demonstrates the potential for use of individual buildings as the spatially disaggregated level-of-activity population, and it provides insight into the activity-based approach and its implication.

APPROACH

To estimate dynamic activity population and validate the results, this study develops a spatial–temporal activity population by using data collected through student travel diary surveys at North Carolina State University (NCSU) in 2001. The travel diary data include student demographic characteristics and travel characteristics (*17, 18*). NCSU GIS mapping provides the locations of campus buildings, activities within them, and floor area (in square feet). The analysis estimates the number of students engaged in each activity (e.g., activity presence) at the level of individual campus building. Analyses show the number of students estimated through student registration records, which include class hour, number of students, and classroom building, to be enrolled in class or study activities. Class or study activities include class lectures, labs, research work, teaching assistant assignments, and similar student academic activities for which students register.

DATA

The NCSU population comprised 26,400 students and about 7,000 employees, including 1,800 faculty members, in 2001. A random 3.2% sample of students representing the total NCSU student population was selected, and those students were asked to complete an activity travel diary for one school day (*17, 18*). The questionnaires collected data on student residential location, student characteristics, location of daily activities, and travel mode to activities.

TABLE 1 Surveyed Students

Classification	Residence Location of Students (%) On Campus	Off Campus	Total
Freshman	159 (90.3)	17 (9.7)	176
Sophomore	130 (79.8)	33 (20.2)	163
Junior	78 (51.0)	75 (49.0)	153
Senior	46 (30.1)	107 (69.9)	153
Graduate	16 (8.1)	182 (91.9)	198
Total	429 (50.9)	414 (49.1)	843 (100.0)

The descriptive summary statistics in Table 1 show that 843 students returned the survey questionnaires. Samples were almost evenly distributed among on-campus and off-campus students. Most freshmen, sophomores, and juniors live on campus, and more of them returned survey questionnaires than did seniors and graduate students, who usually live off campus. Table 2 compares survey characteristics to university enrollment characteristics. Samples by student educational status were appropriately selected to reflect actual enrollment.

As shown in Table 3, the original 843 returned samples were cleaned of nonresponses and were reduced to the surveys of 698 students accomplishing 4,883 daily activities. The questionnaires collected data on student residential location (on-campus, off-campus), student characteristics (gender, age, and status: freshman, sophomore, junior, senior, or graduate student), location of daily activities, and travel mode to activities. Student activities were surveyed by 10 activity classifications: class, study, work or volunteer, drop off or pick up, meals, social or recreation, shop, doctor or other professional, family or personal, and sleep.

As shown in Table 3, the most frequent activity in daily student life is class. Of 4,883 activities, 1,958 (40.1%) are university class or study activities. Meals and social or recreation activities also showed high frequencies. The average number of class activities is the most frequent activity (1.61 per person per day). Students go to the library or a lab to study or conduct research about once daily (1.06). The meals activity is the second most frequent (1.10). Table 3 also shows that students allocate large proportions of their daily time budgets to class and study activities (about 4 h). Drop off or pick up is the shortest activity, whereas school-related activities are the longest overall. Social or recreation also occupies a substantial portion of the

TABLE 2 Comparison of Student Sample to University Population

Classification	Actual (%)	Survey (%)
Freshman	18.6	20.9
Sophomore	16.5	19.3
Junior	17.7	18.1
Senior	21.1	18.1
Graduate	26.0	23.5
Total	99.9	99.9

TABLE 3 Daily Activities

Activity	Frequency[a] (%)	Average Number of Activities[b]	Average Duration (min)[c]	Average Travel Time (min)[d]
Class	1,160 (23.8)	1.61	93.7	11.41
Study	798 (16.3)	1.06	139.9	11.52
Work or volunteer	264 (5.4)	0.40	238.5	15.16
Drop off or pick up someone	152 (3.1)	0.19	14.9	12.03
Meals	847 (17.4)	1.10	50.5	11.48
Social or recreational	562 (11.5)	0.73	86.6	11.84
Shop	219 (4.5)	0.27	21.1	10.80
Doctor or other professional	21 (0.4)	0.03	51.1	20.48
Family or personal activities	519 (10.6)	0.67	59.3	12.02
Sleep	341 (7.0)	0.68	186.8	12.00
Overall weighted average		6.76	93.6	12.00

[a]The total frequency is 4,883 by 698 students after cleaning of nonresponse answers.
[b]Number of activities per person per day.
[c]Average duration per activity (total activity time/total frequency of activity).
[d]Average minutes spent on travel for each activity per day.

average time budget (about 87 min per day). The overall average travel time for activities on campus is 12 min, indicating that the majority of trips are within or near NCSU.

ESTIMATION OF SPATIAL–TEMPORAL ACTIVITY POPULATION

So that activity presence can be estimated, the number of people in each student group must be classified by using university demographic information. Then the number of students by group is applied to the activity schedule. The activity schedule is aggregated by hourly segment and multiplied by the number of students in each class or study activity, resulting in the number of students engaged in each activity in each hour. Next, the activity capacity spatially distributes all students in an hourly time frame. Since this study examines student activity presence at individual buildings, the activity presence defines the number of students in a building for an activity. Hence, activity presence is called activity population. A destination choice model is developed to estimate activity capacity of every building. Finally, students are spatially allocated on the campus.

Traveler Groups

Universities have students, faculty, and staff. However, in this study, activity presence is estimated only for four student groups—on-campus undergraduate, on-campus graduate, off-campus undergraduate, and off-campus graduate—since the activity diary survey is not available for faculty and staff.

Activity Type

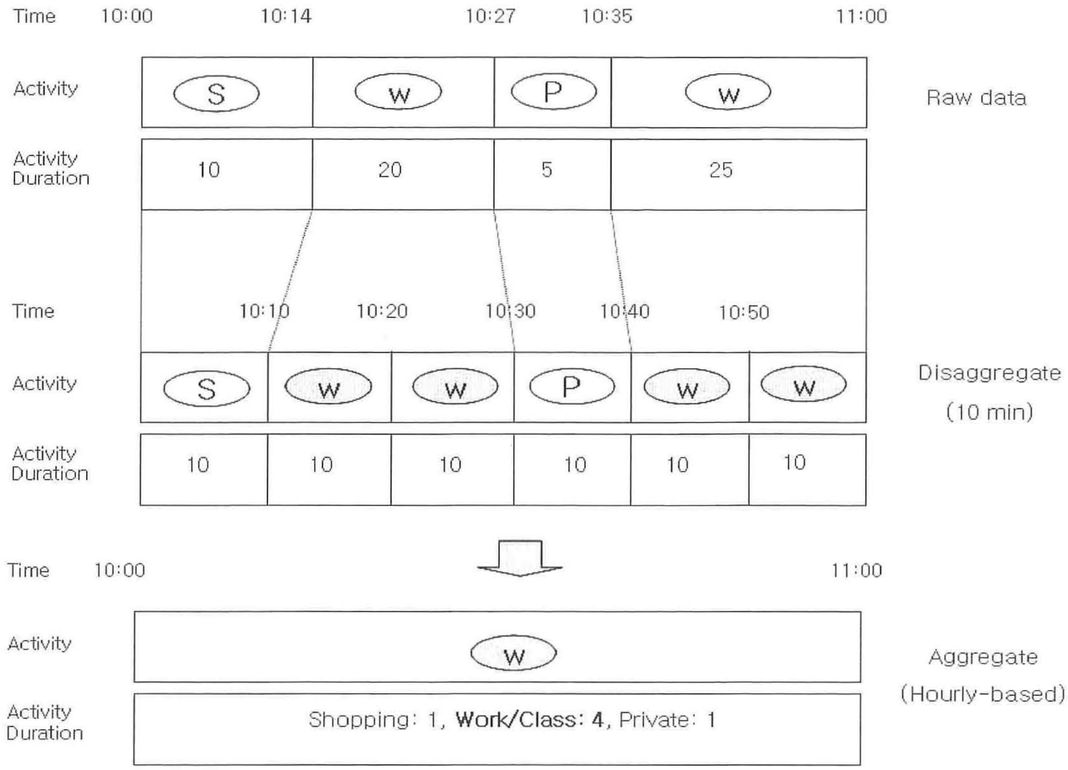
The student activity survey conducted by the NCSU Department of Transportation classifies 10 activities (*14*). In this study, these 10 activities are regrouped into five major activities to simplify model development regarding activity schedule and activity constraint (capacity). The five main activity types are home, class or study, shopping, recreation, and other services.

Synthetic Hourly Activity Schedule

The activity schedule is the sequence of all activities and is obtained from the observed daily activity participation of students as shown by the student activity diary. The observed activity participation in minutes is aggregated into hourly intervals for estimating hourly activity presence. For simplification of the activity schedule, observed activity participation is reclassified by 10-min intervals on the basis of activity duration after adjustment for the minutes left over. Then the major activity is defined as the most frequent activity for every hourly segment. Figure 1 shows the procedure for developing an hourly activity schedule, and Figure 2 shows the hourly activity schedules for student groups. This type of representation is called an activity profile (*10*). This representation facilitates the understanding of daily activity sequences, and it simplifies development of activity-based travel demand models (*19, 20*). Short activities are overlooked when this procedure is used, but the reasonable time segments decrease model complexity.

Activity Capacity

The activity capacity of a location is defined as the value of attractiveness of a building where people can participate in an activity in a time period. In the model, a location with a higher activity capacity can attract more people. The attractiveness of the spatial dimension, where people pursue a certain activity, can be obtained from a building inventory. For this study, the spatial dimension of student activities derives from the NCSU building inventory, which classifies each building by purpose. The information includes building use, building square footage, and number of classroom seats per building, and it helps development of activity capacity in terms of building attractiveness by activity type. A building may have multiple activities.



(a)

Activity	1AM	2AM		10AM	11AM	12PM		11PM	12AM
H	0.962	0.901	....	0.372	0.335	0.325	....	0.911	0.969
C/S	0.015	0.071	....	0.533	0.519	0.480	....	0.034	0.011
S	0.001	0.002	....	0.013	0.013	0.023	....	0.004	0.001
R	0.008	0.007	....	0.055	0.105	0.137	....	0.039	0.012
O	0.013	0.018	....	0.027	0.029	0.036	....	0.012	0.007
SUM	1.000	1.000		1.000	1.000	1.000		1.000	1.000

(b)

FIGURE 1 Development of hourly activity schedule: (*a*) reclassification of activities in hourly time segments and (*b*) aggregate ratios of activity participation.

The activity capacity of a building can be represented by the probability of students' destination choices by facility type. The multinomial logit (MNL) model estimates the probability of destination choice for a campus facility type.

The buildings are differentiated by using the square footage of each to develop individual activity capacities. The activity capacity of an individual building is a function of the probability of the building being chosen and the square footage of the building.

$$C_i^j = \frac{p_k^j \times (S_{i\in k})}{\sum_i \left(p_k^j \times (S_{i\in k})\right)} \qquad (1)$$

where

C_i^j = activity capacity of building i for activity j,
$S_{i\in k}$ = 1,000 ft^2 of building i included in facility type k, and
p_k^j = choice probability for facility type k for activity j (destination choice model).

MNL Destination Choice Model

Choosing a destination for an activity is a form of discrete choice. A random utility-based MNL model for student destination choices is specified. The utility function is defined as the linear form as follows:

$$U_{id} = \alpha_d + \beta_{dP} P_i + \beta_{dT} T_i + \epsilon_{id} \qquad (2)$$

where

U_{id} = utility of facility type d for student i,
α_d = estimable alternative specific constants,
β_d = estimable coefficients,
ϵ_{id} = type i extreme value (Gumbel) distributed random error terms, and
P_i, T_i = vectors of personal and trip information, respectively, for student i.

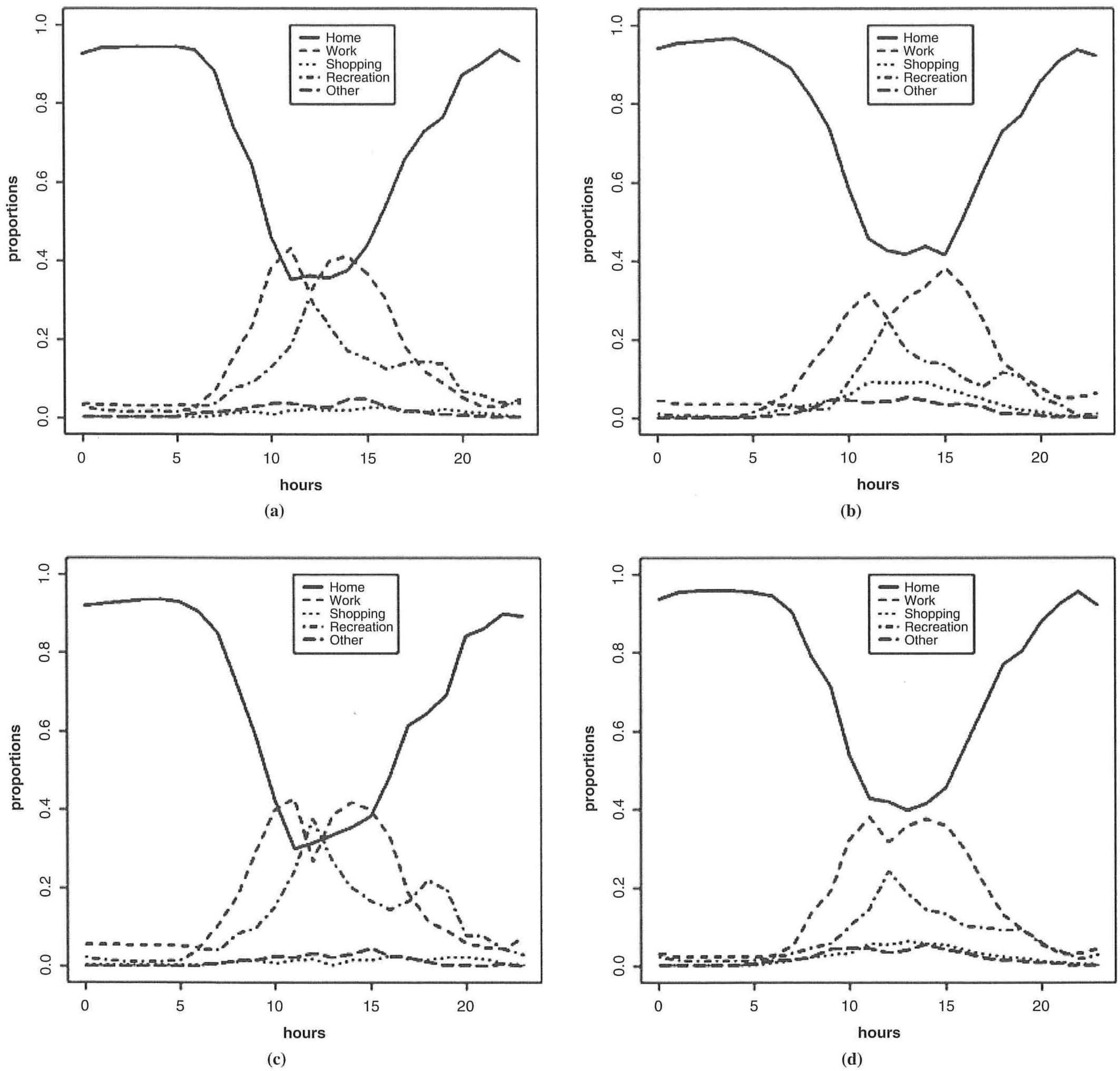


FIGURE 2 Hourly activity schedule by student group: (*a*) undergraduate, (*b*) graduate, (*c*) on-campus, and (*d*) off-campus residents.

If utility maximizing behavior is assumed, this utility leads to the MNL model:

$$P_{id} = \frac{e^{U_{id}}}{\sum_{d \in D_i} e^{U_d}} \tag{3}$$

where

P_{id} = probability of student i choosing facility type d,
U_{id} = utility of facility type d for student i, and
D_i = choice set of all available alternative facility type for student i.

All variables are specific to a particular student. The MNL model shows how the student-specific variables affect destination choice. The unavailable alternative specific information is captured by the alternative specific constant.

MNL Model Estimation

The student survey recorded 2,293 on-campus activities (about 47% of total student activities). The facility inventory shows that all activities do not occur at every facility. Therefore, the destination choice model must be developed by activity type in which different choice sets (facilities) are defined. For example, the recreation activity

does not happen at research facilities. Hence, this facility type is not included in the destination choice model for the recreation activity, whereas all eight facility types are alternatives for both class or study and other.

The facility choice set consists of the eight types:

- Academic or teaching facility,
- Library,
- Student service facility,
- Housing facility,
- Research facility,
- Administration or office facility,
- Athletic facility, and
- Other (student health center, theater, or unclassified).

Table 4 shows the explanatory variables that represent students' personal information and their travel time for individual activities. Personal information includes age, gender, registered credits, education status (undergraduate and graduate), student status (full time and part time), residential status (on and off campus), license to drive, and employment status (full time, part time, volunteer, and unemployed). These variables are tested and used to develop the MNL destination choice model by activity type, such as class or study, recreation, and other. A shopping activity is not considered in the destination choice model because the data set is too small for estimation of parameters.

MNL Model Results

Table 5 is a summary of the model tested globally for the effect of each variable on the outcome variable, with other variables in the model controlled for. All the chi-square statistics are Wald statistics, not likelihood ratio statistics. Each chi-square is a test of the null hypothesis that the explanatory variable has no effect on the outcome variable.

TABLE 4 Explanatory Variables

Variable	Description	Coding
Personal information		
Age	Age in years	Continuous
Credits	Credit hours registered	Continuous
Gender	If the person is male, then 1; otherwise, 0	1: male 0: female
Estatus	Educational status	1: under 0: graduate
Sstatus	Student status	1: full-time 0: part-time
Residence	Residential status	1: on campus 0: off campus
License	Licensed to drive	1: yes 0: no
Empstatus	Employment status	Indicator, 1 if full-time worker Indicator, 1 if part-time worker Indicator, 1 if volunteer Indicator, 1 if unemployed
Travel		
Ttime	Travel time (min)	Continuous

In the destination choice models for the activities class or study and other, there are seven degrees of freedom for each chi-square because each variable has seven coefficients. Thus the null hypothesis is that seven coefficients are zero. The log likelihood ratio equals twice the positive difference between the log likelihoods for the fitted model and the saturated model, and high p-values suggest a good fit. The p-value of 1.000 in all three models ensures that the models fit well (*21*). As shown in Table 5, the chi-square statistics of some variables were not estimated because of the lack of data sets having all possible cases.

TABLE 5 Summary of Maximum Likelihood Analysis of Variance

	Activity					
	Class or Study		Recreation		Other	
Variable	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$	χ^2	$Pr > \chi^2$
Age	21.31	0.0033**	—	—	12.74	0.0786
Credits	15.32	0.0321*	—	—	5.91	0.5499
Gender	5.38	0.6137	35.44	<0.0001**	—	—
Educ. status	—	—	—	—	—	—
Student status	—	—	—	—	—	—
Resident status	18.77	0.0089**	59.87	<0.0001**	—	—
License	—	—	—	—	—	—
Employment	—	—	—	—	—	—
Travel time	16.39	0.0218*	5.82	0.4432	5.43	0.6078
N	1,560		627		97	
df	7		6		7	
LL ratio	1,431.10 (df: 6e3)		140.26 (df: 402)		265.99 (df: 581)	
p-value	1.0000		1.0000		1.0000	

NOTE: '—' indicates that the variable is not estimated. The SAS CATMOD procedure is used in the analysis (*21*).
$^*p < .05$, $^{**}p < .01$.

TABLE 6 MNL Destination Choice Model of Student Activities

		Destination Facility						
Activity	Variable	Academic or Teaching	Library	Service Facility	Housing Facility	Research Facility	Office Facility	Athletic Facility
Class or study[a]	Intercept	8.8461**	5.3247**	—	7.8205*	8.3718*	7.9237*	8.0961*
	Credit	—	—	—	−0.3409*	−0.2524	−0.2842*	—
	Residence	—	—	—	2.8306*	—	—	1.8191
	Age	−0.0753*	—	—	−0.1432	−0.1679*	−0.1342	−0.1972*
	Ttime	—	−0.0883*	—	—	—	—	—
Recreation[b]	Intercept	2.9047**	—	3.1179**	—	—	—	2.5454**
	Gender	—	2.8113*	2.2159**	1.6872*	—	—	2.4609**
	Residence	—	−3.5128*	—	—	—	—	—
Other[c]	Intercept	6.2477	—	21.7019*	17.9458**	—	—	—
	Age	−0.1351	—	−0.8008*	−0.6019**	—	—	—

NOTE: Other facility is defined as a reference variable. Coefficients that were not significant at the 90% level are omitted from the table.
[a]Student service facility is omitted from destination facility because there were no observations.
[b]Research facility is omitted from destination facility because there were no observations.
[c]Library, research, office, and athletic facility are omitted from destination facility because there were no observations.
$*p < .05$, $**p < .01$.

It was found that the explanatory variables age, credit, gender, residential status, and travel time are statistically significant on destination choice behavior, whereas the variables educational status (graduate or undergraduate), student status (part time or full time), license, and employment status are not. This suggests, for example, that younger on-campus students take more credit hours and walk shorter distances, as expected, and that courses taken are related to gender.

Table 6 gives estimation results for the MNL destination choice model of students for the activities class or study, recreation, and other, respectively. This analysis uses the largest value of the dependent variables as a reference. Accordingly, the facility "other" is the reference variable for estimating the parameters for each facility type.

In the destination choice model for the class or study activity, age in years shows that the older the student, the lower the propensity toward most of the facilities except for library and student service facility as compared with the facility other. It was also found that the more credits students are registered for, the lower the propensity toward housing, research, and administration or office. On-campus residents have an increased propensity toward housing and athletic facilities.

From the result of the destination choice model for the recreation activity, gender and residential status explain destination choice. Male students are more likely to choose library, student service, housing, and athletic facilities than are female students as compared with facility choice other. In the destination choice model for the activity choice "other," only age is selected as an explanatory variable because of the lack of data sets for estimating parameters of all variables. The older the student, the lower the propensity toward academic or teaching, student service facility, and housing facility compared to the facility choice other.

Dynamic Activity Population

Activity presence at an individual building is simply calculated from a function of the number of students in each traveler group, their corresponding activity schedules, and the building activity capacity. Hourly activity presence is obtainable since the activity schedule determines the number of activity participations in every hour during a day.

Equation 4 illustrates the function for activity presence at an individual building:

$$Z_{i,j,g}(t) = P_{jg}(t) \cdot N_g \cdot C_{ij} \tag{4}$$

where

$Z_{i,j,g}(t)$ = total number of students in traveler group g engaged in activity type j at time t present at building i,
$P_{jg}(t)$ = probability of participation in given activity j at time t,
N_g = number of students in group g, and
C_{ij} = activity capacity of building i for activity j.

VALIDATION OF ACTIVITY POPULATION

To validate the estimated student activity presence at each building (e.g., activity population), student class schedules for a Tuesday in the spring semester of 2001 were obtained from the NCSU student registration office. The class schedules defined observed activity population data for each building. The registration information included course code, number of students registered, class room, building, and class time.

To create observed activity population data, the number of students registered was sorted by building name and class time. Class time is separated by hour and minute. A Visual Basic program was coded to accumulate the number of students by buildings and hours. To accumulate students, for example, consider that if 28 students are registered for a class running from 9:05 to 10:20 a.m., then the 28 students are counted into both the 9:00 a.m. and the 10:00 a.m. presence (activity population). It is assumed that the registered students attend the class.

Figure 3 is an example of student population in each building in the morning hours (i.e., 9:00 to 10:00 a.m.). The students appear to be evenly distributed at housing and academic or teaching buildings, but similar graphs for other hours show that the number of students at housing buildings decreases as the hours pass and increases at academic or teaching buildings for the class or study activity.

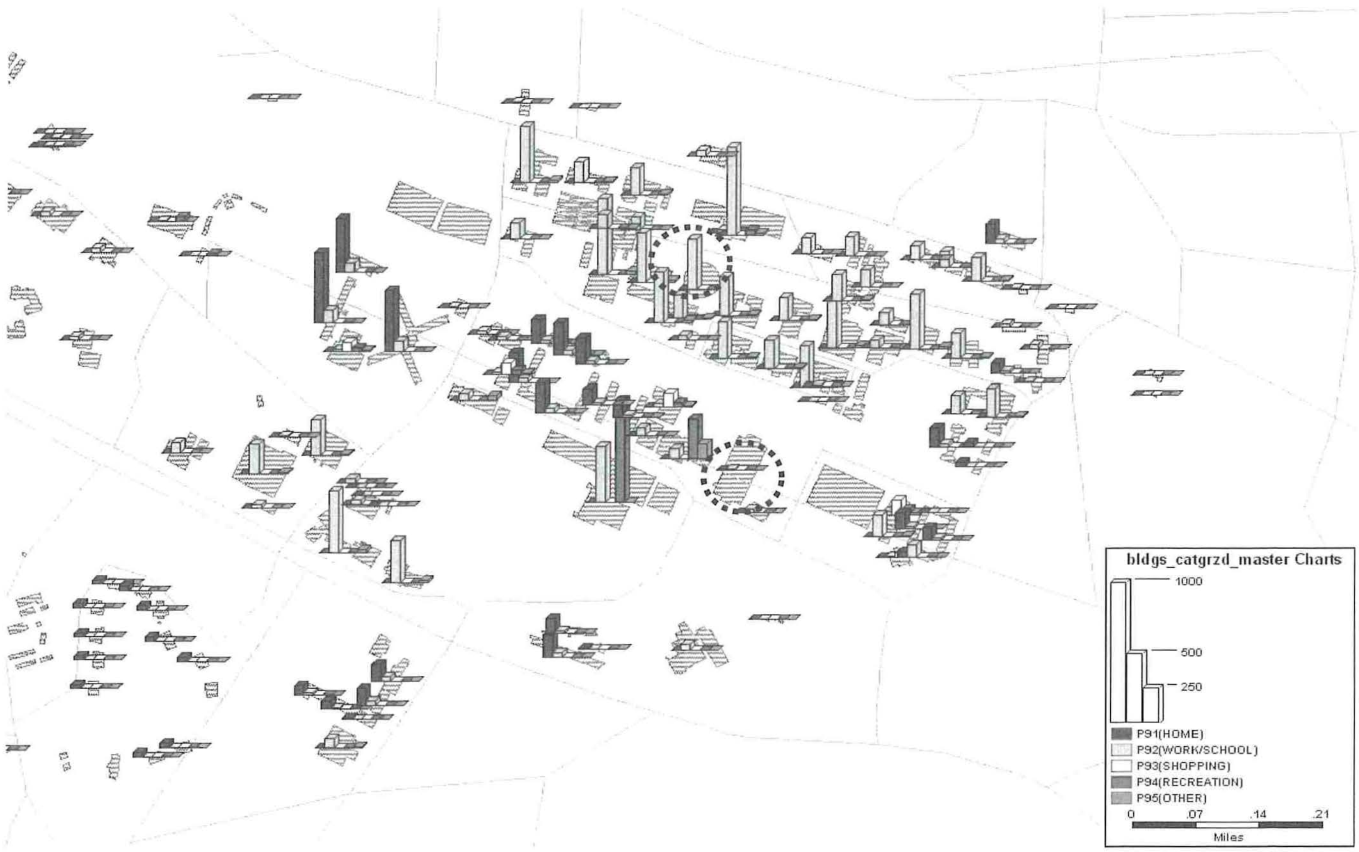


FIGURE 3 Activity population in each building at NCSU, 9:00 to 10:00 a.m.

Conversely, the presence of students is high at housing buildings and low at academic or teaching buildings after 3:00 p.m. Many students were present at student service and athletic buildings in the afternoon for recreational and other activities. This output makes sense because most classes begin before 4:00 p.m. and students often engage in recreational activities after class. The building presence appears to successfully reflect students' daily spatial and temporal activity patterns on campus.

Figure 4 shows the difference between estimated student presence in buildings and number of students recorded. For example, Harrelson Hall is a teaching building (building in the upper dotted circle of Figure 3) and has the highest student presence of any building, which corresponds to observed students registered. Thus, the model appears to successfully replicate the spatial distribution of students in classroom buildings on campus, and the activity capacity in each building based on the destination choice model effectively estimates building presence. Likewise, Reynolds Gym (rectangular building in the lower dotted circle) has the lowest student presence in the morning.

Figure 4*a* compares hourly modeled presence (student population) and registered students for classes in 30 buildings. The model captures the temporal distribution of students during class hours. However, in most hours (Figure 4) a higher number of students was estimated compared to students registered. This difference results from the modeled hourly student presence including students in both the activities class or study and other.

Table 7 is a direct comparison between estimated activity population and number of students registered in a class by building. The overall estimated activity population of 30 buildings is seen to be higher than the observed by as much as 35%. This is mainly because the model considers not only the students in a classroom but also the students doing other activities in the same building. Thus, 25 of 30 buildings are overestimated with various ranges represented by the percent deviation that illustrates the variation of estimates regarding the observations.

To check the accuracy of model estimates, the R^2 is calculated and reveals that the overall modeled building presence explains approximately 81% of the variation based on 30 buildings during 13 hourly periods (from 8:00 a.m. to 8:00 p.m.), and the modeled hourly building presence explains about 67% of the variation. In both cases, R^2 is low but is acceptable for prototype models, because the models were not calibrated and the observed values include only the students in class.

SUMMARY AND CONCLUSIONS

This study introduced the spatial–temporal activity-based approach for estimating and validating activity population at the individual building level for a case study at NCSU. The spatial and temporal class or study activity population of students was estimated by development of the hourly activity schedule and activity capacity in each building on the NCSU campus. Data from student registration records were obtained to validate the class or study activity population. The major findings from the empirical validation of model estimations for hourly activity population at NCSU are summarized as follows:

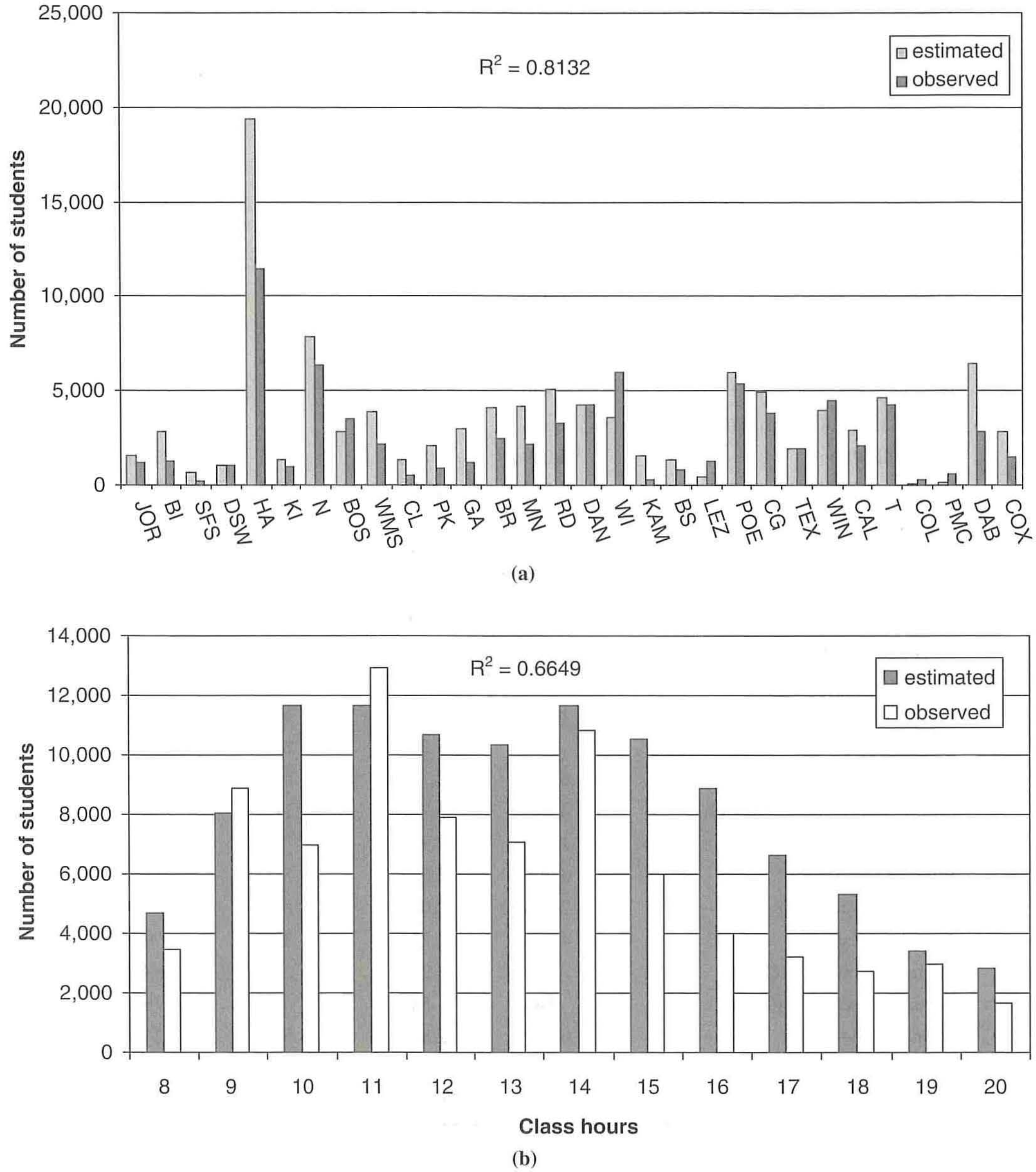


FIGURE 4 Comparison of estimated and observed student presence: (*a*) 30 buildings during class hours (8:00 a.m. to 8:00 p.m.) and (*b*) hourly modeled and observed student presence.

- The estimated number of students in each building (activity population) was directly compared to the actual number of students registered for classes in a building. The comparison illustrated that the overall estimated activity population replicated actual students relatively well, although the estimated population included all students, both in class and not.
- Although the activity capacity in each building based on the destination choice model effectively estimated of activity population, other factors affecting the choice of building need to be investigated and included in the model system for better replication of observed numbers. These factors include building development density, ingress and egress options, and transportation access options.

As a prototype model, this study provides a solid basis for suggesting that building-level spatial–temporal models are feasible and useful. The relative simplicity and accuracy of the models suggest that the additional level of detail can overcome many problems in aggregate models. Although the methodologies used in the analysis are not novel, the procedure of creating an activity schedule and activity distribution is straightforward, easily understood, and useful to practitioners, especially university planners. Hence, this study is a meaningful reference to practitioners supporting modeling activities at the building level. Tasks for future research include the testing of scenarios of changes in building ingress and egress and sidewalk capacities and predicting shifts from auto-oriented to pedestrian- and transit-oriented travel as building density increases. The last task includes the challenge of integrating a land use model at the microscopic level.

Several other issues were not addressed because of research scope and data limits. First, this study considered student activity-travel

TABLE 7 Estimated and Registered Populations

Building Name	Estimates	Students Registered	Absolute Difference	Deviation (%)
Jordan Hall (JOR)	1,570	1,229	341	27.8
Biltmore Hall (BI)	2,840	1,304	1,536	117.7
Schaub Food Science Building (SFS)	642	234	408	174.5
D.S. Weaver Labs (DSW)	1,066	1,011	55	5.4
Harrelson Hall (HA)	19,414	11,438	7,976	69.7
Kilgore Hall (KI)	1,320	966	354	36.7
Nelson Hall (N)	7,864	6,338	1,526	24.1
Bostian Hall (BOS)	2,821	3,514	693	−19.7
Williams Hall (WMS)	3,879	2,162	1,717	79.4
David Clark Laboratories (CL)	1,313	553	760	137.2
Polk Hall (PK)	2,059	896	1,163	129.7
Gardner Hall (GA)	2,982	1,190	1,792	150.5
Broughton Hall (BR)	4,134	2,442	1,692	69.3
Mann Hall (MN)	4,217	2,157	2,060	95.5
Riddick Engineering Labs (RD)	5,083	3,266	1,817	55.6
Daniels Hall (DAN)	4,267	4,302	35	−0.8
Withers Hall (WI)	3,589	5,997	2,408	−40.1
Kamphoefner Hall (KAM)	1,599	334	1,265	378.3
Brooks Hall (BS)	1,333	814	519	63.8
Leazar Hall (LEZ)	442	1,261	819	−65.0
Poe Hall (POE)	6,023	5,390	633	11.7
Carmichael Gymnasium (CG)	4,929	3,782	1,147	30.3
Textile Building (TEX)	1,968	1,950	18	0.9
Winston Hall (WIN)	3,991	4,503	512	−11.4
Caldwell Hall (CAL)	2,948	2,092	856	40.9
Tompkins Hall (T)	4,616	4,236	380	9.0
Reynolds Coliseum (COL)	44	273	229	−83.9
Price Music Center (PMC)	144	584	440	−75.4
Dabney Hall (DAB)	6,439	2,865	3,574	124.8
Cox Hall (COX)	2,808	1,508	1,300	86.2
Total	106,342	78,590	27,752	35.3

NOTE: The estimation includes students both in class and not in class. The observations count only students in class. Percent deviation is calculated as

$$\%\text{deviation} = \frac{(\text{obs}_i - \text{est}_i)}{\text{obs}_i} \times 100$$

behavior from a 1-day travel survey. The exploratory analysis should be extended to more days, other populations (faculty, staff, and visitors), other universities or organizations, and communities that have available housing, registration, and employment records. Second, model evaluation should be carried out in a regional travel demand model either by incorporating the activity-based approach within conventional four-step modeling or by developing an activity-based model for an entire region. It would be interesting to see how the activity-based approach improves model accuracy at an aggregated level of model development, since the activity-based approach is expected to provide better estimates of trips for a regional level. Finally, further comparisons are needed between conventional and various activity-based modeling approaches to provide additional insight into activity-based approaches for planning agencies and model practitioners responding to a range of scenarios and policies.

ACKNOWLEDGMENTS

The authors acknowledge Slade McCalip of the NCSU Department of Transportation for help with the student activity travel survey data and Leta F. Huntsinger of the Triangle Regional Model Service Bureau for supporting this study.

REFERENCES

1. Algers, S., J. Eliasson, and L. Mattsson. Activity-Based Model Development to Support Transport Planning in the Stockholm Region. Presented at 5th Workshop of TLE Network, Nynashamn, Sweden, 2001.
2. *Activity-Based Modeling System for Travel Demand Forecasting*. RDC, Inc. Sept. 1, 1995. http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/amos/amos.pdf/. Accessed Oct. 4, 2006.

3. Rossi, T. *TMIP Activity and Tour-Based Model Seminar.* Jan. 9. 2005. http://www.trb-forecasting.org/TRBWorkshop-158_Rossi.pdf/. Accessed Nov. 1, 2006.
4. Pendyala, R., and C. R. Bhat. Validation and Assessment of Activity-Based Travel Demand Modeling System. Presented at Innovations in Travel Demand Modeling Conference, Austin, Tex., 2006.
5. Hägerstrand, T. What About People in Regional Science? *Regional Science Association,* Vol. 24, 1970, pp. 7–21.
6. Chapin, F. S. *Human Activity Patterns in the City.* John Wiley and Sons, New York, 1974.
7. Kitamura, R., S. Fujii, T. Yamamoto, and A. Kikuchi. Application of PCATS/DEBNets to Regional Planning and Policy Analysis: Micro-simulation Studies for the Cities of Osaka and Kyoto, Japan. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2001.
8. Arentze, T., and H. Timmermans. *ALBATROSS: A Learning Based Transportation Oriented Simulation System.* European Institute of Retailing and Service Studies, Eindhoven, Netherlands, 2000.
9. *A System of Activity-Based Models for Portland, Oregon.* Travel Model Improvement Program, U.S. Department of Transportation, 1998.
10. Kuhnau, J. *Activity-Based Travel Demand Modeling Using Spatial and Temporal Models in the Urban Transportation Planning System.* MS thesis. Pennsylvania State University, University Park, 2001.
11. Goulias, K. G. Transportation Planning and Modeling. In *Handbook of Transportation Engineering* (M. Kutz, ed.), McGraw–Hill, New York, 2003.
12. Eom, J. *Empirical Spatial-Temporal Activity-Based Data Analysis: Spatial–Temporal Activity Population and Travel Demand of University.* VDM Verlag, New York, 2009.
13. Alam, S. B., and Goulias, K. G. Dynamic Emergency Evacuation Management System Using Geographic Information System and Spatio-temporal Models of Behavior. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1660,* TRB, National Research Council, Washington, D.C., 1999, pp. 92–99.
14. Kwan, M. Interactive Geovisualization of Activity-Travel Patterns Using 3-D Geographical Information Systems: A Methodological Exploration with a Large Data Set. *Transportation Research Part C,* Vol. 8, No. 1–6, 2000, pp. 185–203.
15. Buliung, R. N., and P. S. Kanroglou. A GIS Toolkit for Exploring Geographics of Household Activity/Travel Behavior. *Journal of Transport Geography,* Vol. 14, No. 1, 2006, pp. 35–51.
16. Beckx, C. B., L. I. Panis, T. Arentze, D. Janssens, R. Torfs, S. Broekx, and G. Wets. A Dynamic Activity-Based Population Modeling Approach to Evaluate Exposure to Air Pollution: Methods and Application to a Dutch Urban Area. *Journal of Environmental Impact Assessment Review,* Vol. 29, No. 3, 2009, pp. 179–185.
17. *Analysis of NCSU Travel Demand Survey.* MAB, Inc., Raleigh, N.C., 2004.
18. Eom, J., J. R. Stone, and S. K. Ghosh. Daily Activity Patterns of University Students. ASCE *Journal of Urban Planning and Development,* Vol. 135, No. 4, 2009, pp. 141–149.
19. Ma, J. *An Activity-Based Approach and Micro-simulated Travel Forecasting System: A Pragmatic Synthetic Scheduling Approach.* PhD thesis. Pennsylvania State University, University Park, 1997.
20. Bhat, C., and F. S. Koppelman. A Retrospective and Prospective Survey of Time-Use Research. *Transportation,* Vol. 26, No. 2, 1999, pp. 119–139.
21. Allison, P. D. Logistic Regression Using the SAS System: Theory and Application. SAS Institute, Cary, N.C., 1999.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Changes in Travel Behavior in Response to Weather Conditions

Do Type of Weather and Trip Purpose Matter?

Mario Cools, Elke Moons, Lieve Creemers, and Geert Wets

Weather can influence travel demand, traffic flow, and traffic safety. A hypothesis—the type of weather determined the likelihood of a change in travel behavior, and changes in travel behavior because of weather conditions depended on trip purpose—was assayed. A stated adaptation study was conducted in Flanders (the Dutch-speaking region of Belgium). A survey, completed by 586 respondents, was administered both on the Internet and as a traditional paper-and-pencil questionnaire. To ensure optimal correspondence between the survey sample composition and the Flemish population, observations in the sample were weighted. To test the main hypotheses, Pearson chi-square independence tests were performed. Results from both the descriptive analysis and the independence tests confirm that the type of weather matters and that changes in travel behavior in response to these weather conditions are highly dependent on trip purpose. This dependence of behavioral adjustments on trip purpose provides policy makers with a deeper understanding of how weather conditions affect traffic. Further generalizations of the findings are possible by shifting the scope toward revealed travel behavior. Triangulation of both stated and revealed travel behavior on the one hand and traffic intensities on the other hand is a key challenge for further research.

A deeper understanding of how various weather conditions affect traffic is essential for policy makers. This is stressed by policy issues that often are related to adverse weather events, such as increased fuel consumption, economic losses due to traffic delays, and higher traffic counts. At the network level, adverse weather events increase uncertainty in system performance, resulting, for instance, in a network capacity reduction ranging from 10% to 20% in heavy rain (*1*). Maze et al. reported that weather events affect three predominant traffic variables: travel demand, traffic safety, and the traffic flow relationship (*2*). This paper focuses on the effect of weather on travel demand.

Weather can influence travel demand in various ways, including diversions of trips to other modes or other paths and cancellation of trips (*2*). Day-to-day weather conditions such as fog and precipitation can decrease travel demand, for instance, when drivers postpone or cancel discretionary activities (e.g., leisure activities), but can also have an increasing effect when travel modes are shifted from slow modes (walking, cycling) to motorized vehicles (*3*). Mode changes, changes in departure time, and diversions to an alternate route were reported by Khattak and De Palma as the most prevalent behavioral adaptations (*4*). Bos indicated that in the Netherlands, heavy rain reduces the number of cyclists (*5*), whereas mild winters and warm summers increase bicycle use. Van Berkum et al. noted that the reduction in bicycle use during heavy rain is accompanied by a modal shift from bicycle to car (for either driver or passenger) (*6*). A similar result was found by Nankervis (*7*), who examined the effect of both (short-term) weather conditions and (long-term) seasonal variation patterns on bicycle commuting patterns among students in the temperate climate of Melbourne, Australia: cycle commuting was affected by long-term, climatic conditions as well as daily weather conditions. According to Guo et al., temperature, rain, snow, and wind all influence transit ridership of the Chicago Transit Authority: good weather increases ridership, whereas bad weather has a diminishing effect (*8*). Guo et al. stressed that vehicle running and dwell times, as well as cost of operation, also are affected by weather (*8*). In Brussels, Belgium, however, the transit agency reported higher levels of transit ridership during adverse weather (*4*).

The main objectives for this paper are to test the hypothesis that the type of weather influences the likelihood of a change in travel behavior (e.g., assessing whether more people change their transport mode during snow than during periods of fog) and to assay whether changes in travel behavior due to weather conditions are dependent on trip purpose (e.g., examining whether because of snowy weather, more people cancel leisure and shopping trips than school- or work-related trips). To this end, a stated adaptation study was conducted in Flanders, the Dutch-speaking region of Belgium.

METHODOLOGY

Stated Adaptation Approach

The data needed to address the main research questions were collected by means of a stated adaptation experiment. Various descriptions of stated adaptation experiments can be found in the literature (*9, 10*). In this paper, stated adaptation experiments are regarded as an alternative to the more widely used stated preference and choice experiments. The main difference between stated adaptation and stated preference and choice experiments is the task posed to respondents. In stated preference experiments, respondents are asked to indicate their preference to sequentially presented attribute profiles. In stated choice experiments, respondents are shown choice sets of two or more attribute profiles and are asked to choose the profile they like best (or, alternatively, to allocate some fixed budget among the profile).

Transportation Research Institute, Hasselt University, Wetenschapspark 5, Bus 6, BE-3590 Diepenbeek, Belgium. Corresponding author: G. Wets, geert.wets@uhasselt.be.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 22–28.
DOI: 10.3141/2157-03

In stated adaptation experiments, respondents are asked to indicate if and how they would change their behavior considering experimentally varied attribute profiles, typically representing scenarios.

A stated adaptation survey, administered both on the Internet (86.7%) and via a traditional paper-and-pencil questionnaire (13.3%), was completed by 586 respondents. This dual-mode administration remedied the sample bias that is introduced when only an Internet-based data collection is conducted; studies have demonstrated that some socioeconomic classes, such as older-age and lower-education groups, may be more reluctant to use computer-assisted instruments for data collection (*11, 12*). In total, 90 behavioral adaptations in response to different weather conditions were queried; the frequencies of five changes in travel behavior in response to six weather conditions were surveyed, and this observation was repeated for three types of trips.

Weather Conditions

The following weather conditions were considered: cold temperature, defined as temperatures below freezing; warm temperature, defined as temperatures above 28°C (82.4°F); snow or freezing rain; heavy rain or thunderstorm (abbreviated as rain); fog; and storm or heavy wind. Cools et al. reported that these weather conditions had a significant effect on daily traffic intensities measured on Belgian highways (*13*). Therefore, the decision was made to analyze the effect of these weather conditions on underlying travel behavior. How frequently these weather events occur is demonstrated by the various weather-related measures in Table 1 (*14*). Flanders has a moderate maritime climate.

TABLE 1 Weather Parameters, Uccle, Belgium (*14*)

Parameter	2007	2008	Normal[a]
Average wind speed (m/s)	3.3	3.4	3.7
Sunshine duration (h)	1,472.0	1,449.0	1,554.0
Average temperature (°C)	11.5	10.9	9.7
Average maximum temperature (°C)	15.3	14.6	13.8
Average minimum temperature (°C)	7.8	7.2	6.7
Absolute maximum temperature (°C)	30.9	31.0	31.7
Absolute minimum temperature (°C)	−6.8	−6.1	−8.9
Number of freezing days (min. < 0°C)	27.0	37.0	47.0
Number of wintry days (max. < 0°C)	1.0	0.0	8.0
Number of summery days (max. ≥ 25°C)	23.0	25.0	25.0
Number of heat wave days (max. ≥ 30°C)	2.0	1.0	3.0
Average relative atmospheric humidity (%)	80.0	77.0	81.0
Total precipitation (mm)	879.5	861.5	804.8
Number of days with measurable precipitation (≥ 0,1 mm)	204.0	209.0	207.0
Number of days with thunderstorm	94.0	95.0	94.0

[a]Normal: long-term meteorological average (1971–2000).

Changes in Travel Behavior

The stated adaptation questionnaire was split into three parts, corresponding to the three types of trips that were considered for the analysis. These three trip types correspond to the categories of most commonly performed trips according to the Flemish travel behavior survey (*15*): commuting (work, school), shopping, and leisure trips. Equivalent questions were asked in each part: for a certain behavioral change, the respondents were to indicate how often (never, in 1% to 25% of the cases, in 26% to 50% of the cases, or in more than 50% of the cases) they make a certain change in travel behavior for each of the six weather conditions. The following changes in travel behavior were queried: (*a*) change in transport mode, (*b*) change in timing of the trip [postponement (advancing) of the trip to a later (earlier) time on the same day], (*c*) change in the location where the activity (work or school, shopping, or leisure) will be performed, (*d*) elimination of the trip by skipping the activity (trip cancellation), and (*e*) change in the route of the trip. As an illustration of the questionnaire style, Figure 1 shows a question concerning the postponement or advancing

Do you postpone or advance your work/school-related trip to a later/earlier moment the same day due to any of the following weather conditions?

Mark the answer that corresponds mostly to your situation. Only ***one*** *answer is possible for* ***each weather condition****.*

	No, never	Yes, occasionally (<25% of the cases)	Yes, sometimes (<50% of the cases)	Yes, usually (>50% of the cases)
Cold temperature	O	O	O	O
Snow/freezing rain	O	O	O	O
Heavy rain/thunderstorm	O	O	O	O
Fog	O	O	O	O
Warm temperature	O	O	O	O
Storm/heavy wind	O	O	O	O

FIGURE 1 Stated adaptation question about postponement or advancing of work- or school-related trips.

of work- or school-related trips to a later or earlier moment in the same day.

Statistical Analyses

So that there is an optimal correspondence between the survey sample composition and the Flemish population, the observations in the sample are weighted. The weights were calculated by matching the marginal distributions of the sample with the marginal distributions of the population. Age, gender, and civil state were the basis for this matching process. Recall that the main objectives for this paper are to test the hypothesis that the type of weather determines the likelihood of a change in travel behavior and to assay whether changes in travel behavior because of weather conditions are associated with the type of trip. To test these hypotheses, independence tests are performed.

To test independence (this is the null hypothesis) between two multinomial (categorical) variables, one could use the Pearson statistic Q_p, which is defined by the following equation:

$$Q_p = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{\left(n_{ij}-\hat{\mu}_{ij}\right)^2}{\hat{\mu}_{ij}}$$

where n_{ij} is the observed frequency in cell (i, j), calculated by multiplying the observed chance by the sample size, and $\hat{\mu}_{ij}$ is the expected frequency for table cell (i, j). When the row and column variables are independent, Q_p has an asymptotic chi-square distribution with $(k-1)(l-1)$ degrees of freedom (*16*).

A criticism of the Pearson statistic is that is does not give a meaningful description of the degree of dependence (or strength of association). Cramer's contingency coefficient, often referred to as Cramer's *V*, is a method for interpreting the strength of association and is calculated with the following formula:

$$V = \sqrt{\frac{Q_p}{N(m-1)}}$$

where

Q_p = Pearson chi-square statistic defined earlier,
N = total sample size, and
m = minimum of number of rows and number of columns in the contingency table.

Basically, Cramer's *V* scales the chi-square statistic Q_p to a value between 0 (no association) and 1 (maximum association). Note that Cramer's *V* has the desirable property of scale invariance, that is, if the sample size increases, the value of Cramer's *V* does not change so long as values in the contingency table change the same relative to each other (*17*).

DESCRIPTIVE ANALYSIS OF CHANGES IN TRAVEL BEHAVIOR

Changes in Commuting Trips

For the commuting (work, school) trips, the percentages of respondents making a certain travel behavior change are given in Table 2. When different weather conditions are compared, it is clear that snow has the largest effect on commuting trips. Time-of-day decisions (postponing the trip to a later moment) especially are common practice: more than one person in two appears to postpone his or her trip in the presence of snow. Route taken is changed by almost half the respondents. This major impact of snow on travel behavior is revealed on the network. For example, Hanbali and Kuemmel found traffic volume reductions on highways away from the major urban centers in the United States ranging from 7% to 56%, depending on the intensity of the snowfall (*18*).

Extreme temperatures (both cold and warm) appear to have the least impact on commuting behavior, whereas storms, fog, and heavy rain appear to have an effect mainly on the timing of the trip: people appear to postpone their trips until more favorable weather conditions apply.

When the focus turns to behavioral changes, it is immediately clear that the work or school location is the least frequently changed.

TABLE 2 Changes in Work and School Trips

Change	Frequency (%)	Cold (%)	Snow (%)	Rain (%)	Fog (%)	Warm (%)	Storm (%)
Mode change	Never	93.8	75.8	84.8	94.6	81.6	86.8
	1–25	4.4	14.6	7.9	3.7	10.5	8.1
	26–50	0.9	2.6	1.4	0.1	4.4	0.9
	>50	0.9	7.0	5.9	1.6	3.5	4.2
Time-of-day change	Never	89.5	47.8	70.3	74.0	94.4	74.9
	1–25	6.0	23.7	17.0	13.7	2.8	14.9
	26–50	2.5	9.2	6.9	6.9	1.5	4.7
	>50	2.0	19.3	5.8	5.4	1.3	5.5
Location change	Never	96.6	86.6	94.4	97.5	97.0	93.3
	1–25	2.2	8.4	3.3	1.3	2.0	4.1
	26–50	0.6	3.0	1.0	0.5	0.8	1.1
	>50	0.6	2.0	1.3	0.7	0.2	1.5
Trip cancellation	Never	96.2	75.4	93.8	95.3	89.0	92.6
	1–25	3.4	19.4	5.0	4.3	10.1	6.1
	26–50	0.4	4.1	0.2	0.4	0.9	0.7
	>50	0.0	1.1	1.0	0.0	0.0	0.6
Route change	Never	90.5	56.4	85.0	85.4	96.4	87.1
	1–25	6.3	26.7	9.9	10.0	2.4	8.4
	26–50	1.8	9.8	2.5	1.5	0.9	2.7
	>50	1.4	7.1	2.6	3.1	0.3	1.8

Obviously, the locations of work and school sites are relatively fixed. Nonetheless, telecommuting alternatives, satellite offices, and e-learning are opportunities that make location change feasible. The most prevalent changes in commuting behavior are changes in the timing of the trip and changes in the route chosen. A possible reason for this is that people try to avoid traffic jams by diverging their paths and changing the departure times of their trips.

Changes in Shopping Trips

The percentages of respondents making a certain travel behavior change for shopping trips are displayed in Table 3. Similar to commuting trips, the most marked finding from the comparison of weather conditions is that snow has the largest effect on shopping trips. Time-of-day changes (trip postponements) and trip cancellations especially are standard: about 70% of the respondents postpone their shopping trips, and the same percentage cancels their shopping trips.

The effects of heavy rain and heavy winds or storms are striking: about 60% of the respondents postpone and around 50% cancel their shopping trips during periods of heavy rain, and about 50% of the respondents postpone their shopping trips during stormy periods, whereas 45% cancel their shopping trips.

A comparison of extreme temperatures provides the insight that more people change their transport mode for shopping trips during warm temperatures (above 28°C) than during cold temperatures (below freezing). A possible explanation is that people are more interested in using slow modes (walking, cycling) during highly favorable weather conditions. This is in line with the results of Bos (*5*), who found an increase in bicycle use during warm summers.

The most prevalent changes in shopping-related travel behavior are trip postponements and trip cancellations. Extreme weather conditions appear to cause serious changes in the activities people want to perform. When the overall results of shopping trips are compared with commuting trips, a considerably larger percentage of people change their shopping-related travel behavior than change their commuting behavior. This change occurs because it is much easier to postpone or cancel shopping activities than work- or school-related activities, which is also observed on the Flemish highway network (*13*).

Changes in Leisure Trips

For the final category of trips that were considered, namely, leisure trips, the percentages of respondents making certain travel behavior changes are shown in Table 4. Yet again, snowy weather has the largest impact. Similar to shopping trips, trip postponements and trip cancellations are the most frequent changes in travel behavior: about 65% of the respondents postpone their leisure trips, and the same number of respondents cancel their leisure trips.

Heavy rain and heavy wind or storms also clearly influence leisure-related travel behavior: about 45% of the respondents postpone and a similar percentage cancel their leisure trips during periods of heavy rain or heavy wind. The effect of extreme temperature on leisure trips is similar to the effect on shopping trips: more people alter their transport modes for leisure trips during warm temperatures than during cold temperatures. The resemblances between shopping and leisure trips are further underlined when the most prevalent changes in leisure trips are discussed: trip postponements and trip cancellations are the most frequently performed changes in leisure-related travel behavior. There is homogeneity between behavioral changes concerning leisure trips and shopping trips, because both leisure and shopping activities are nonobligatory activities, which are more flexible (more easy to postpone, advance, or cancel) than are obligatory activities such as school and work.

STATISTICAL ANALYSIS OF CHANGES IN TRAVEL BEHAVIOR

The descriptive results in the previous section clearly indicate that changes in travel behavior in response to weather conditions are dependent on the type of weather condition. Moreover, the results suggested that behavioral changes are strongly dependent on the

TABLE 3 Changes in Shopping Trips

Change	Frequency (%)	Cold (%)	Snow (%)	Rain (%)	Fog (%)	Warm (%)	Storm (%)
Mode change	Never	91.5	78.2	85.6	91.9	79.7	86.8
	1–25	5.2	11.2	6.0	4.4	10.2	6.5
	26–50	1.4	3.4	2.2	0.8	4.9	1.6
	>50	1.9	7.2	6.2	2.9	5.2	5.1
Time-of-day change	Never	80.2	29.4	41.8	59.9	80.0	47.7
	1–25	13.1	28.2	24.1	19.2	13.0	22.8
	26–50	3.9	16.9	13.6	11.4	4.2	13.7
	>50	2.8	25.5	20.5	9.5	2.8	15.8
Location change	Never	86.8	54.0	68.4	72.2	83.7	69.3
	1–25	7.4	20.6	12.6	11.9	10.5	13.7
	26–50	2.8	9.4	10.7	8.8	2.6	10.0
	>50	3.0	16.0	8.3	7.1	3.2	7.0
Trip cancellation	Never	86.7	31.9	48.4	64.4	82.6	55.0
	1–25	7.1	33.7	29.3	20.4	13.3	23.3
	26–50	3.0	14.5	11.6	8.8	2.7	11.6
	>50	3.2	19.9	10.7	6.4	1.4	10.1
Route change	Never	93.1	58.8	81.7	80.6	93.3	81.7
	1–25	4.5	23.2	11.0	11.3	4.7	10.7
	26–50	1.4	10.3	3.7	4.8	0.5	4.6
	>50	1.0	7.7	3.6	3.3	1.5	3.0

TABLE 4 Changes in Leisure Trips

Change	Frequency (%)	Cold (%)	Snow (%)	Rain (%)	Fog (%)	Warm (%)	Storm (%)
Mode change	Never	89.9	74.4	83.9	87.3	77.3	85.6
	1–25	7.7	13.5	8.9	8.1	11.7	8.7
	26–50	1.2	3.8	3.1	3.5	6.4	3.0
	>50	1.2	8.3	4.1	1.1	4.6	2.7
Time-of-day change	Never	85.3	35.1	54.3	61.8	85.3	58.6
	1–25	10.5	30.9	26.1	21.3	11.5	20.1
	26–50	2.0	15.0	12.7	9.2	2.0	13.0
	>50	2.2	19.0	6.9	7.7	1.2	8.3
Location change	Never	83.3	70.9	75.1	81.5	83.9	74.1
	1–25	9.9	14.1	11.3	9.3	10.0	13.1
	26–50	2.8	6.5	6.3	5.3	3.3	6.5
	>50	4.0	8.5	7.3	3.9	2.8	6.3
Trip cancellation	Never	79.3	35.6	56.1	66.1	82.2	55.3
	1–25	14.4	34.0	24.2	20.2	13.9	23.5
	26–50	4.1	13.8	9.6	8.0	3.0	12.1
	>50	2.2	16.6	10.1	5.7	0.9	9.1
Route change	Never	92.8	55.1	76.4	78.6	94.3	76.9
	1–25	4.4	24.4	13.9	13.5	3.6	12.4
	26–50	2.1	11.9	5.9	4.5	1.2	6.9
	>50	0.7	8.6	3.8	3.4	0.9	3.8

underlying activity. In this section, these two hypotheses are formally tested by using Pearson chi-square independence tests. A statistical analysis is provided of the hypothesis that the type of weather determines the likelihood of a change in travel behavior. An elaboration on the test of the hypothesis that changes in travel behavior due to weather conditions are dependent on trip purpose is made. Multiple testing is accounted for by lowering the significance level in a Bonferroni-like approach.

Dependence of Changes in Travel Behavior on Type of Weather

For each activity (trip purpose), the dependency between change in travel behavior and type of weather was formally tested. Table 5 gives the chi-square values, degrees of freedom, and corresponding significance levels of the different tests. First, for each activity, the dependency between all travel behavior changes and weather conditions was tested. Then the dependencies of specific travel behavior changes and weather conditions were assessed.

It can be concluded from Table 5 that all behavioral changes highly depend on the type of weather. (The null hypothesis of independence is rejected for all behavioral changes with a p-value smaller than .001.) Similar to the preliminary conclusions drawn from the descriptive results, for work- and school-related trips trip postponement (time-of-day change) and route changes are the strongest depending on the weather type; for shopping and leisure trips, the relationship is the most significant (higher chi-squared value, same degrees of freedom) for trip postponements and trip cancellations. Although highly significant (p-value smaller than .001), the interdependence

TABLE 5 Dependence of Changes on Weather

Trip Purpose	Behavioral Change	χ^2	df	Signif.	Cramer's V
Work or School	All changes	1,185.75	95	***	0.125
	Mode change	138.71	15	***	0.123
	Time-of-day change	409.05	15	***	0.212
	Location change	81.12	15	***	0.094
	Trip cancellation[a]	174.79	5	***	0.240
	Route change	362.56	15	***	0.199
Shopping	All	1,728.89	95	***	0.142
	Mode change	92.24	15	***	0.095
	Time-of-day change	542.97	15	***	0.230
	Location change	235.69	15	***	0.152
	Trip cancellation	555.65	15	***	0.233
	Route change	302.34	15	***	0.172
Leisure	All	1,456.24	95	***	0.130
	Mode change	107.92	15	***	0.102
	Time-of-day change	522.45	15	***	0.224
	Location change	62.85	15	***	0.078
	Trip cancellation	405.26	15	***	0.197
	Route change	357.76	15	***	0.185

[a]Estimated using reduced answer possibilities (yes/no).
***p-value < .001.

between changes in travel behavior and type of weather was smallest for location and mode changes.

Recall that the number of degrees of freedom is calculated by multiplying the number of rows minus one by the number of columns minus one. For the dependence of specific travel behavior changes on weather conditions, the independence test followed a chi-square distribution with 15 degrees of freedom: four frequencies (the number of people who would never change their behavior and, respectively, the ones who change their behavior in 1% to 25%, 26% to 50%, and more than 50% of the cases) minus one multiplied by six weather conditions minus one. Since the underlying assumption of the independence test (minimum 80% of the cells expected counts should be at least equal to 5) was not fulfilled for the hypothesis test that assessed the relationship between trip cancellations of commuting trips and weather conditions, an alternative independence test was tabulated by combining the three categories of people that change their behavior. As a result, the number of degrees of freedom for this test was smaller than that for the other test, as shown in Table 5.

Dependence of Changes in Travel Behavior on Trip Purpose

To test the dependence of changes in travel behavior on activity type (trip purpose), independence tests are performed on an aggregate level (aggregation over all travel behavior changes). Table 6 displays the chi-square values, degrees of freedom, and corresponding significance levels of these tests.

In line with tests that assess the dependence of changes in travel behavior on type of weather, the dependence of changes in travel behavior on trip purpose also confirm the preliminary conclusions drawn from the descriptive results: the extent to which people adapt their travel behavior is strongly dependent (all p-values smaller than .001) on trip purpose. This dependence appears to be the largest for periods of heavy rain, snow, and heavy wind. For extreme temperatures, this dependency appears to be smaller (lower chi-square value and same number of degrees of freedom) yet still highly significant.

To further investigate the dependence of changes in travel behavior on trip purpose, a more-detailed analysis is performed: for all six weather conditions, the dependence of the specific behavioral changes on trip purpose was investigated. Various conclusions could be drawn from this disaggregate analysis. First, for all weather conditions, time-of-day changes (trip postponements), location changes, and trip cancellations significantly depended on trip purpose (p-values all smaller than .01 and for location changes all smaller than 0.001).

TABLE 6 Dependence of Changes on Trip Purpose, Aggregate Level

Weather	Behavioral Change	χ^2	df	Signif.	Cramer's V
All types	All changes	2,180.35	238	***	0.148
Cold	All changes	165.69	38	***	0.100
Snow	All changes	473.46	38	***	0.169
Rain	All changes	550.80	38	***	0.183
Fog	All changes	382.66	38	***	0.152
Warm	All changes	144.80	38	***	0.094
Storm	All changes	462.94	38	***	0.167

***p-value < .001.

TABLE 7 Dependence of Changes on Trip Purpose, Disaggregate Level

Weather	Behavioral Change	χ^2	df	Signif.	Cramer's V
Cold	Mode change	9.03	6	NS	0.052
	Time-of-day change	21.24	6	**	0.080
	Location change	50.88	6	***	0.124
	Trip cancellation	79.41	6	***	0.155
	Route change	5.12	6	NS	0.039
Snow	Mode change	5.07	6	NS	0.039
	Time-of-day change	49.55	6	***	0.122
	Location change	143.46	6	***	0.208
	Trip cancellation	271.33	6	***	0.287
	Route change	4.06	6	NS	0.035
Rain	Mode change	9.93	6	NS	0.055
	Time-of-day change	129.11	6	***	0.198
	Location change	120.21	6	***	0.191
	Trip cancellation	275.88	6	***	0.289
	Route change	15.68	6	*	0.069
Fog	Mode change	42.05	6	***	0.113
	Time-of-day change	30.06	6	***	0.095
	Location change	126.67	6	***	0.196
	Trip cancellation	170.08	6	***	0.227
	Route change	13.80	6	*	0.065
Warm	Mode change	5.41	6	NS	0.040
	Time-of-day change	54.24	6	***	0.128
	Location change	58.03	6	***	0.133
	Trip cancellation[a]	11.59	2	**	0.059
	Route change[a]	5.15	2	NS	0.039
Storm	Mode change	13.15	6	*	0.063
	Time-of-day change	97.85	6	***	0.172
	Location change	104.97	6	***	0.178
	Trip cancellation	225.54	6	***	0.261
	Route change	21.42	6	**	0.081

[a]Estimated using reduced answer possibilities (yes/no).
*p-value < .05, **p-value < .01, ***p-value < .001, NS = not significant.

In addition, all behavioral changes in response to fog and heavy wind or storm were statistically significantly depending on the trip purpose (p-values all smaller than .05).

A thorough look at the effects of cold and warm weather, as well as snow, shows that the extent to which people change their mode or route in response to these weather conditions does not depend on trip purpose. Furthermore, Table 7 shows that for all weather conditions (except warm weather), the highest dependency of behavioral changes on trip purpose was found for trip cancellations. A possible explanation for this contrast with warm-weather conditions is that all other weather conditions are considered to be unfavorable, whereas high temperatures may be considered as favorable, at least for some part of the population.

CONCLUSIONS

In this study, the hypothesis of dependence of changes in travel behavior on type of weather and the hypothesis of dependence of changes in travel behavior on trip purpose (activity type) were formally tested. Results from both the descriptive analysis and the Pearson chi-square independence tests confirm that the type of weather matters and that the changes in travel behavior in response to weather conditions are highly dependent on trip purpose.

Whereas most papers in the international literature focus on traffic safety and traffic flows, this paper contributes to the literature by

looking at the underlying travel behavior by means of a multifaceted stated adaptation approach. The clear dependence of behavioral adjustments on activities (trip purposes) provides policy makers with a deeper understanding of how weather conditions affect traffic. The value of this contribution is stressed by weather-related policy issues, such as increased fuel consumption, economic losses due to traffic delays, and higher traffic counts.

The findings in this paper are consonant with international literature and provide a solid basis for further analysis of weather-related policy measures, such as an examination of whether extreme weather conditions cause last-minute changes in travel mode, or an assessment of whether high-quality bus shelters make a difference in last-minute mode changes. Further generalizations of the findings are possible by shifting the scope toward revealed travel behavior and by breaking down modal changes to different transport modes. Triangulation of both stated and revealed travel behavior on the one hand and traffic intensities on the other hand is a key challenge for further research.

REFERENCES

1. De Palma, A., and D. Rochat. Understanding Individual Travel Decisions: Results from a Commuters Survey in Geneva. *Transportation,* Vol. 26, No. 3, 1999, pp. 263–281.
2. Maze, T. H., M. Agarwal, and G. Burchett. Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1948,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 170–176.
3. Hranac, R., E. Sterzin, D. Krechmer, H. Rahka, and M. Farzaneh. *Empirical Studies on Traffic Flow in Inclement Weather.* FHWA-HOP-07-073. FHWA, U.S. Department of Transportation, 2006.
4. Khattak, A. J., and A. De Palma. The Impact of Adverse Weather Conditions on the Propensity to Change Travel Decisions: A Survey of Brussels Commuters. *Transportation Research Part A,* Vol. 31, No. 3, 1997, 181–203.
5. Bos, J. M. J. *In All Kinds of Weather: Road Safety Effects of Periods of Extreme Weather* (in Dutch). Stichting Wetenschappelijk Onderzoek Verkeersveiligheid, Leidschendam, Netherlands, 2001.
6. Van Berkum, E., W. Weijermars, and A. Hagens. The Impact of Weather on Urban Travel Demand in the Netherlands. *Proc., EWGT2006 International Joint Conferences.* Bari, Italy, 2006, pp. 245–252.
7. Nankervis, M. The Effect of Weather and Climate on Bicycle Commuting. *Transportation Research Part A,* Vol. 33, No. 6, 1999, pp. 417–431.
8. Guo, Z., N. H. M. Wilson, and A. Rahbee. Impact of Weather on Transit Ridership in Chicago, Illinois. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2034,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 3–10.
9. Lee-Gosselin, M. E. H. Scope and Potential of Interactive Stated Response Data Collection Methods. In *Conference Proceedings 10: Household Travel Surveys: New Concepts and Research Needs,* TRB, National Research Council, Washington, D.C., 1996, pp. 115–133.
10. Faivre D'Arcier, B., O. Andan, and C. Raux. Stated Adaptation Surveys and Choice Process: Some Methodological Issues. *Transportation,* Vol. 25, No. 2, 1998, pp. 169–185.
11. Couper, M. P., A. Kapteyn, M. Schonlau, and J. Winter. Noncoverage and Nonresponse in an Internet Survey. *Social Science Research,* Vol. 36, No. 1, 2007, pp. 131–148.
12. Hayslett, M. M., and B. M. Wildemuth. Pixels or Pencils? The Relative Effectiveness of Web-Based Versus Paper Surveys. *Library and Information Science Research,* Vol. 26, No. 1, 2004, pp. 73–93.
13. Cools, M., E. Moons, and G. Wets. Assessing the Impact of Weather on Traffic Intensity. *Weather, Climate, and Society,* Vol. 2, No. 1, 2010, pp. 60–68.
14. Royal Meteorological Institute. *Climatic Overview of 2008* (in Dutch). http://www.meteo.be/meteo/view/nl/2827848-2008.html. Accessed July 15, 2009.
15. Zwerts, E., and E. Nuyts. *Travel Behavior Research Flanders (January 2000–January 2001), Part 3B: Analysis of the Persons Questionnaire* (in Dutch). Department of Architecture, Provinciale Hogeschool Limburg, Diepenbeek, Belgium, 2004.
16. Agresti, A. *Categorical Data Analysis,* 2nd ed. John Wiley and Sons, Hoboken, N.J., 2002.
17. Sheskin, D. J. *Handbook of Parametric and Nonparametric Statistical Procedures,* 2nd ed. Chapman and Hall, Boca Raton, Fla., 2000.
18. Hanbali, R. M., and D. A. Kuemmel. Traffic Volume Reductions Due to Winter Storm Conditions. In *Transportation Research Record 1387,* TRB, National Research Council, Washington, D.C., 1993, pp. 159–164.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Assessing the Impact of Public Holidays on Travel Time Expenditure

Differentiation by Trip Motive

Mario Cools, Elke Moons, and Geert Wets

The impact of public holidays on the underlying reasons for travel behavior, namely, the activities people perform and the trips made, is seldom investigated. Therefore, the effect of holidays on travel time expenditure in Flanders, differentiated by trip motive, is examined. The data used for the analysis stem from a household travel survey carried out in 2000. The zero-inflated Poisson regression approach is used; it explicitly takes into account the inherent contrast between travelers and nontravelers. The zero-inflated Poisson regression models yield findings that are harmonious with international literature: sociodemographic variables, temporal effects, and transportation preferences contribute significantly to unraveling the variability of travel behavior. In particular, it is shown that the effect of public holidays on daily travel behavior cannot be ignored. Triangulation of quantitative and qualitative techniques is a solid basis for insight into the underpinnings of travel behavior.

The importance of a thorough examination of the effect of public holidays on travel time expenditure was underlined by Liu and Sharma (*1*) and Cools et al. (*2, 3*), who stressed the need to incorporate holiday effects in travel behavior models. First, public holidays can influence both the demand for activities (e.g., on regular days, the demand for work activities is much larger than it is for periods during which most people plan their holidays) and the supply of activity opportunities in space and time (e.g., operating hours of amusement parks are often prolonged during holiday periods). Second, holidays can affect the supply of available transport options (e.g., during summer holidays, extra trains and planes are scheduled to transfer people to popular holiday destinations). Finally, holidays can influence the supply of infrastructure and their associated management systems (e.g., during the summer holiday period, police often enforce driving in groups to limit traffic congestion).

The literature on holiday effects largely concerns two areas: the effect of holidays on traffic counts (*4, 5*) and on traffic safety (*6, 7*). The impact on underlying reasons for travel behavior, namely, the activities people perform and the trips made, is seldom investigated. Therefore, this study discusses the effect of public holidays on trips made, with a focus on attribute travel time.

It is important to differentiate travel time expenditure by trip motive. First, commuting (which is defined as work- and school-related trips), although the main reason for performing trips, accounts for only 26.8% of all trips (*8*). Thus, a focus on commuting trips would neglect almost three-quarters of all trips reported. Likewise, concentration of the analysis on shopping (defined as both daily and nondaily shopping, 20.5% of all trips) or leisure trips (14.2% of all trips) is to be avoided.

Differentiation by trip motive can trigger a refinement of the underlying relationships between travel behavior and explanatory factors. With a division of the travel time expenditure into subparts dependent on trip motive, more-complex relationships can be implicitly modeled: differentiation makes it feasible to incorporate explanatory factors that have an increasing or decreasing effect on a particular subpart and that have an opposite effect, a substitution effect, or no effect at all on other subparts.

OVERVIEW OF DATA

Correspondence of Sample to Population

The data used for the analysis stem from a household travel survey in Flanders that was carried out in 2000 (*8*). This survey was done to investigate the travel behavior of people living in the Flanders area. Through stratified clustered sampling, 3,028 households were queried about their travel behavior. All household members older than 6 (7,625 persons total) were asked to report the trips they made during a particular day, which yielded information on about 21,031 trips.

To guarantee an optimal correspondence between the survey sample composition and the population, the observations in the sample were weighted. The weights were calculated by matching the marginal distributions of the sample with the marginal distributions of the population. Age, gender, and civil state were the basis for this matching process.

Dependent Variables: Travel Time Expenditure by Purpose

The daily travel time expenditure for each trip motive was calculated by adding the time spent on trips related to the specific

Transportation Research Institute, Hasselt University, Wetenschapspark 5, Bus 6, BE-3590 Diepenbeek, Belgium. Corresponding author: G. Wets, geert.wets@uhasselt.be.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 29–37.
DOI: 10.3141/2157-04

motive. Both trips to the activity locations and trips back home were considered.

Explanatory Variables

Temporal Effects

The first category of explanatory variables used in the analysis is temporal effects; the first temporal effect considered is the day-of-week effect. Agarwal showed that there exists a significant difference between travel behavior on a weekday and travel behavior on a weekend day (*9*). This difference is further revealed by Sall and Bhat (*10*) and Schwanen (*11*) by demonstration of a significant day-of-week effect. In the analysis, the day-of-week effect is represented by a categorical variable with seven categories, the first category corresponding to a Monday, the last to a Sunday.

The focus of this study is on the second temporal effect, namely, the holiday effect. To evaluate the significance of public holidays on daily commuting time, a special holiday variable is created, consisting of three categories: normal days, holidays, and summer holidays. The following holidays are taken into account: Christmas vacation, spring half-term, Easter vacation, Labor Day, Ascension Day, Whit Sunday, Whit Monday, vacation of the construction industry (three weeks, starting the second Monday of July), Our Blessed Lady Ascension, fall break (including All Saints' Day and All Souls' Day), and Remembrance Day. Note that for all these holidays, the adjacent weekends were considered to be a holiday, too. For holidays occurring on a Tuesday or a Thursday, the Monday and weekend before and the Friday and weekend after, respectively, were also defined as a holiday, because often people have those days off and thus have a leave of several days, which may be used for a short holiday (*2*). The days in July and August not in the preceding holiday list were labeled as summer holidays.

Sociodemographics

In addition to temporal effects, sociodemographic variables were considered in the analysis, as they are commonly used in models that predict travel time (*12–14*). The following variables are considered for the analyses presented in this paper: age, gender, employment status, living conditions, and degree of urbanization.

Transportation Preferences

The final group of variables used for the analysis is frequency of use of various transport modes. The following modes were considered: use of scheduled-service bus and tramway, categorized as people who never, occasionally (a few times a year or month), and frequently (weekly or more often) use this service; use of the railroad system (same categorization as for bus); daily use of a bicycle (dummy variable that equals 1 if the respondent uses the bicycle daily); and daily use of a motorcycle (cf. daily bicycle use). Reports concerning the Flemish travel survey reveal that more than half the respondents never use buses or trams (*8*). The use of trains appears to be slightly more popular. In addition to various transportation uses, possession of a driving license is considered for the analysis.

DESCRIPTIVE ANALYSIS

Dependent Variables: Travel Time Expenditure by Purpose

The distribution categories for travel time expenditure, differentiated by trip motive, are given in Table 1. The table shows that commuting is the most-performed travel activity (it has the smallest percentage of no travel), followed by shopping and leisure trips. In addition to the overall means, the means excluding zeros are tabulated. There are large discrepancies between these two measures of central tendency, suggesting the need for a modeling approach that explicitly takes into account the excess of zeros.

Explanatory Variables

Temporal Effects

Mean travel time expenditures according to trip motive for the categories of the temporal effects are given in Table 2. Time spent on commuting is considerably lower during holidays compared to regular days, and travel time expenditure on leisure trips is portentously higher during holiday periods. Less-pronounced differences are seen for shopping trips. There is a large discrepancy between weekdays and weekend days for commuting travel times, and a lesser discrepancy for leisure travel times. Shopping-related travel times appear to peak on Saturdays.

Sociodemographics

An exploratory analysis of the most dominant sociodemographical variables, shown in Table 2, reveals that the daily time spent on commuting increases with age, reaches its maximum at age category 35 to 44, and declines after people reach retirement age. The daily commuting time appears to be higher for males than for females, and the professionally active population spends more time commuting than the inactive population. Table 2 provides preliminary insight into the travel time spent on shopping trips: travel time increases with age, and females spend more travel time on shopping trips than males. For employment status, the inactive population spends more travel time on shopping then does the active. The overall picture for travel

TABLE 1 Travel Time Expenditure, Differentiated by Trip Motive

Descriptive Measure	Commuting (%)	Shopping (%)	Leisure (%)
Distribution category			
No travel	62.1	70.1	78.3
1–10 min	4.7	8.3	4.8
11–20 min	7.6	8.0	5.0
21–30 min	6.6	5.2	2.9
31–40 min	4.5	2.7	2.0
41–50 min	3.1	1.5	1.5
51–100 min	7.5	3.4	3.2
>100 min	4.0	1.0	2.3
Central tendency			
Mean (with 0s)	18.5 min	8.9 min	10.7 min
Mean (without 0s)	48.9 min	29.8 min	49.5 min

TABLE 2 Mean Travel Time Expenditure by Trip Motive

Explanatory Variable	Commuting (min)	Shopping (min)	Leisure (min)
Holiday			
No holiday	21.9	8.6	8.3
Holiday	11.3	9.6	15.7
Summer holiday	12.8	9.4	16.6
Day of week			
Monday	29.6	7.6	9.4
Tuesday	31.4	6.8	8.6
Wednesday	24.2	8.8	5.7
Thursday	28.5	8.7	7.7
Friday	25.7	8.0	11.6
Saturday	4.4	15.2	16.9
Sunday	3.0	7.0	24.3
Age, years			
6–12	10.8	5.3	14.4
13–15	21.0	3.8	12.1
16–24	27.1	6.1	13.4
25–34	27.7	8.8	9.9
35–44	28.0	8.8	10.7
45–54	24.0	10.7	10.4
55–64	10.0	12.4	13.4
65+	1.0	10.1	6.5
Gender			
Male	24.0	7.5	12.5
Female	13.4	10.2	9.1
Employment status			
Housekeeping	0.6	15.3	9.6
Unemployed	1.7	15.5	6.3
Retired	0.6	10.3	8.6
Disabled	1.1	9.1	8.8
Pupil, student	18.6	4.8	13.8
Worker	30.5	7.8	8.1
Employee	31.7	10.0	11.8
Executive	42.0	8.6	11.9
Liberal profession	15.5	5.2	18.5
Self-employed	20.3	6.1	11.6
Overall	18.5	8.9	10.7

time spent on leisure trips is less striking; however, travel time spent on leisure trips is higher for males than for females and is remarkably lower for the oldest age category (65+).

METHODOLOGY

Zero-Inflated Poisson Regression

The main modeling approach used for the analysis is zero-inflated Poisson (ZIP) regression. This modeling framework uses a ZIP distribution to deal with the excess of zeros. The approach assumes that the population consists of two types of individuals. The first type gives a Poisson-distributed count, which may be zero, whereas the second type always gives a zero count. This assumption can be supported by the inherent contrast between travelers and nontravelers, which could explain discrepancies between the means incorporating and disregarding zeros. The choice for the ZIP regression approach implies that the three types of travel time expenditures will be treated as count variables. The comparison of a linear regression and a Poisson regression model for predicting commuting times revealed that the Poisson regression model explained more of the variability in travel time expenditure on commuting (*15*). Therefore, accommodation of a Poisson model that takes into account the inherent contrast between travelers and nontravelers is a defensible approach. Although travel time expenditures are traditionally analyzed with Tobit models and hazard-based duration models (*16*), in this paper the suitability of the ZIP regression as an alternative modeling framework is illustrated.

The zero-inflated Poisson distribution has two parameters: the mean of the Poisson distribution λ_i and the proportion of individuals of the second type (the nontravelers), ω_i. Formally, the zero-inflated Poisson distribution can be represented as follows (*17*):

$$\Pr(Y_i = k) = \begin{cases} \omega_i + (1-\omega_i)e^{-\lambda_i} & \text{for } k = 0 \\ (1-\omega_i)\dfrac{e^{-\lambda_i} - \lambda_i^k}{k!} & \text{for } k > 0 \end{cases}$$

where both the probability ω_i and the mean number λ_i depend on covariates. For the covariate matrices **B** and **G** of the models discussed in this paper, the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ satisfy the following equations:

$$\begin{cases} \log(\lambda) = \mathbf{B}\boldsymbol{\beta} \\ \operatorname{logit}(\omega) = \log\left(\dfrac{\omega}{(1-\omega)}\right) = \mathbf{G}\boldsymbol{\gamma} \end{cases}$$

Estimates for the unknown parameters are obtained by maximizing the log likelihood by using a ridge-stabilized Newton–Raphson algorithm (*18*). The log likelihood function for the zero-inflated Poisson distribution is given by

$$\sum_{i=1}^{n} l_i$$

where

$$l_i = \begin{cases} w_i \log\left[\omega_i + (1-\omega_i)e^{-\lambda}\right] & k = 0 \\ w_i\left[\log(1-\omega_i) + k\log(\lambda_i) - \lambda_i - \log(k!)\right] & k > 0 \end{cases}$$

where n is the number of observations and where w_i are the weights calculated by matching the marginal distributions of the sample with the marginal distributions of the population. Note that in contrast to the ordinary Poisson regression model, no scale parameter can be included in the ZIP regression model to accommodate for overdispersion (*18*).

Model Performance Assessment

To assess the appropriateness of the zero-inflated Poisson distribution, the van den Broek score test for testing zero inflation relative to a Poisson distribution (*19*) will be performed. The statistic is based on a comparison of the actual zeros to those predicted by the model:

$$S = \frac{\left[\sum_{i=1}^{n}\left\{\dfrac{I(y_i = 0) - p_{0i}}{p_{0i}}\right\}\right]^2}{\sum_{i=1}^{n}\left\{\dfrac{1-p_{0i}}{p_{0i}}\right\} - n\cdot\bar{y}}$$

where

S = score,
$I(y_i = 0)$ = indicator function that is 1 if a given observation equals zero and is zero otherwise,
p_{0i} = probability of a zero for observation i under the null distribution (regular Poisson distribution),
$\bar{y}$ = mean of the observations, and
n = number of observations.

The probability is allowed to vary by observation. The S score is assumed to follow a chi-squared distribution with one degree of freedom.

Two model-selection criteria that balance model fit against model parsimony are tabulated. The first measure is the corrected Akaike information criterion (AICC), given by

$$\mathrm{AICC} = -2\mathrm{LL} + 2p\frac{n}{n-p-1}$$

where

p = number of parameters estimated in the model,
n = number of observations, and
LL = log likelihood evaluated at the value of the estimated parameters (*18*).

A second, similar measure is the Bayesian information criterion (BIC), defined by

$$\mathrm{BIC} = -2\mathrm{LL} + p\log(n)$$

AICC and BIC are useful criteria for selecting among different models, with smaller values representing better models. Simonoff offers an extensive discussion about the use of AICC and BIC with generalized linear models (*20*).

RESULTS

Overall Results

The variables used in the final zero-inflated Poisson regression models and their likelihood ratio statistics are given in Table 3. The table shows that all three categories of variables (sociodemographic variables, temporal effects, and transportation preferences) contribute significantly to the unraveling of daily travel time. The final models also take into account interdependencies between trips, as the travel time spent on a certain type of trip significantly influences the likelihood of performing other trips and the travel time of these other trips, especially in the case of commuting trips.

Concerning the covariates in the Poisson regression part of the model, the holiday effect, the day-of-week effect, age, gender, employment status, degree of urbanization, use of buses and trams, use of trains, and the indicator of making other types of trips play a significant role in all three models. For explanatory variables in the zero-inflation part of the model, the day-of-week effect, gender, employment status, and time spent on other types of trips are significant covariates in all three models. The degree of urbanization did not contribute significantly to any of the zero-inflation parts. Except for the covariate driving license, all other explanatory

TABLE 3 Likelihood Ratio Statistics for ZIP Regression Models

		Commuting		Shopping		Leisure	
Selected Variable	df[a]	Chi^2	p-Value	Chi^2	p-Value	Chi^2	p-Value
Model predicting λ							
Holiday	2	84.2	<.001	7.7	.021	2,052.4	<.001
Day of week	6	401.0	<.001	95.8	<.001	615.2	<.001
Age, years	7	1,291.7	<.001	330.2	<.001	664.9	<.001
Gender	1	15.4	<.001	46.2	<.001	70.3	<.001
Interaction age * gender	7	1,388.2	<.001	238.7	<.001	896.9	<.001
Employment status	9	845.1	<.001	383.5	<.001	1,534.4	<.001
Living conditions	4	—	—	223.6	<.001	1,164.2	<.001
Degree of urbanization	3	120.0	<.001	219.9	<.001	863.0	<.001
Uses of bus or tram	2	931.8	<.001	497.6	<.001	117.2	<.001
Uses of trains	2	3,272.5	<.001	29.1	<.001	27.6	<.001
Daily use of motorcycle	1	86.0	<.001	—	—	99.3	<.001
Daily use of bicycle	1	341.4	<.001	—	—	—	—
Driving license	1	30.0	<.001	—	—	211.4	<.001
Other type trips made	1	1,911.8	<.001	927.4	<.001	7,125.9	<.001
Model predicting ω							
Holiday	2	218.3	<.001	—	—	6.4	.041
Day of week	6	909.1	<.001	136.5	<.001	204.8	<.001
Age, years	7	—	—	17.5	.014	15.5	.030
Gender	1	11.7	<.001	22.3	<.001	30.5	<.001
Employment status	9	1,400.6	<.001	68.1	<.001	29.2	<.001
Living conditions	4	—	—	—	—	14.3	.006
Driving license	1	—	—	17.9	<.001	7.8	.005
Time spent on other type trips	1	357.4	<.001	89.4	<.001	60.3	<.001
Performance measure							
AICC		75,489		46,880		70,572	
BIC		75,908		47,344		71,088	
Score-test (p-value)		<.001		<.001		<.001	

[a]df: degrees of freedom; indicates that the variables are not included in the final model.

variables representing transportation preferences were left out of the zero-inflation part to prevent convergence problems in the estimation procedure.

For the three different types of trips considered, the best model was chosen each time by using the AICC and BIC criteria. The corresponding values for these criteria are displayed in the lower part of Table 3. The necessity of using a zero-inflated Poisson model rather than a regular Poisson model is formally tested by using the van den Broek score test. For all three models, the corresponding p-value is smaller than 0.001, indicating that a zero-inflated Poisson distribution seriously outperforms a regular Poisson distribution for these models.

Commuting Time

The parameter estimates of the zero-inflated Poisson regression model for predicting travel time expenditure on commuting are shown in Table 4. A distinction must be made between the parameters in the model predicting the mean response λ and the parameters for estimating the probability of the zero-inflation ω. The parameters of the Poisson part of the zero-inflated Poisson model (λ) should be interpreted as multiplicative effects. Take as an example the parameter estimates for daily users of a motorcycle. The multiplicative effect of being a daily motorcycle user instead of a nondaily motorcycle user can then be calculated in the following way: $\exp(-0.441 - 0) = \exp(-0.441) = 0.643$. This means that the commuting time of daily motorcycle users is only 64.3% of the commuting of nondaily motorcycle users, given that they share the same characteristics for all other variables. The parameters of the logistic part of the zero-inflated Poisson model (ω) could be seen as log odds ratio multiplicative effects. Take as an example the parameter of the time spent on other types of trips: an increase of 1 min travel time spent on other types of trips has as a consequence that the odds of noncommuting (a zero for travel time expenditure on commuting trips) equals $\exp(0.016) = 1.02$ times the odds of commuting.

When certain covariates are used for modeling both the mean response λ and the probability of zero-inflation ω, the assessment of the overall effect is not straightforward. When both parameters support the same conclusion, the multiplicative effect of the Poisson parameter is elevated by the zero-inflation parameter. Take as an example the comparison between regular days and days within the summer holiday

TABLE 4 ZIP Regression Parameter Estimates for Travel Time Expenditure on Commuting

Parameter	Est.	SE	Parameter	Est.	SE	Parameter	Est.	SE
Poisson Model λ								
Intercept	3.699	0.020	Gender & age, years			Use of buses or trams		
Holiday			Male, 6–12	−0.429	0.031	Frequently	0.337	0.011
Regular day	0.000		Male, 13–15	−0.658	0.033	Occasionally	0.125	0.008
Holiday	−0.061	0.009	Male, 16–24	−0.494	0.021	Never	0.000	
Summer holiday	−0.120	0.015	Male, 25–34	−0.193	0.019	Use of trains		
Day of week			Male, 35–44	0.000		Frequently	0.531	0.012
Monday	0.000		Male, 45–54	0.093	0.022	Occasionally	−0.038	0.008
Tuesday	−0.063	0.010	Male, 55–64	−0.256	0.034	Never	0.000	
Wednesday	−0.087	0.010	Male, 65+	9.004	3.747	Daily use of motorcycle		
Thursday	−0.037	0.010	Employment status			Yes	−0.441	0.044
Friday	−0.046	0.010	Housekeeping	0.039	0.079	No	0.000	
Saturday	−0.230	0.018	Unemployed	−0.340	0.059	Daily use of bicycle		
Sunday	−0.203	0.024	Retired	−0.012	0.042	Yes	−0.143	0.008
Gender			Disabled	−0.456	0.093	No	0.000	
Male	0.456	0.014	Pupil, student	−0.117	0.018	Driving license		
Female	0.000		Worker	0.000		Yes	0.081	0.014
Age, years			Employee	0.136	0.009	No	0.000	
6–12	−0.367	0.028	Executive	0.200	0.011	Other type trips made		
13–15	0.275	0.028	Liberal profession	−0.142	0.038	Yes	−0.293	0.007
16–24	0.247	0.019	Self-employed	−0.134	0.017	No	0.000	
25–34	0.086	0.016	Degree of urbanization					
35–44	0.000		Metropolitan area	−0.155	0.014			
45–54	−0.072	0.019	Urban area	0.018	0.008			
55–64	0.188	0.030	Suburban area	−0.049	0.012			
65+	−8.891	3.747	Rural area	0.000				
Zero Inflation ω								
Intercept	−2.032	0.162	Day of week			Employment status		
Holiday			Friday	0.060	0.151	Retired	4.703	0.306
Regular day	0.000		Saturday	2.573	0.163	Disabled	4.235	0.528
Public holiday	1.245	0.103	Sunday	3.205	0.203	Pupil, student	0.407	0.128
Summer holiday	1.473	0.154	Gender			Worker	0.000	
Day of week			Male	−0.290	0.090	Employee	−0.088	0.130
Monday	0.000		Female	0.000		Executive	−0.383	0.179
Tuesday	−0.107	0.160	Employment status			Liberal profession	1.031	0.368
Wednesday	0.116	0.150	Housekeeping	4.970	0.500	Self-employed	1.043	0.191
Thursday	0.077	0.152	Unemployed	3.785	0.362	Time spent on other trips	0.016	0.001

period: the parameters of the Poisson parameter indicate that the average commuting time on a regular day is 1.13 (= exp(0 + 0.120)) times the commuting time during a day within the summer holiday period, and this effect is enlarged by the zero-inflation part indication that the odds of commuting are 4.33 for regular days compared to summer holidays. However, when both parameters support opposite effects, the assessment of the overall effect remains inclusive. Consider the difference between Saturdays and Sundays: whereas the Poisson parameters indicate that the commuting time on Sundays is 1.03 (= exp(−0.203 + 0.230)) times the commuting time on Saturdays, the zero-inflation parameters indicate that the odds of commuting on a Saturday versus a Sunday are 1.87 (= exp(3.205 − 2.573)).

Examination of temporal effects shows that the traditional organization of modern society into 5-day workweeks predominates the travel time expenditure on commuting: the likelihood of commuting and the average time spent on commuting are considerably larger during weekdays than during weekend days. This finding is consistent with the results reported by Bhat and Misra (*21*) and Sall and Bhat (*10*), who indicated the importance of incorporating day-of-week effects to account for variability in travel times. Furthermore, travel time expenditure is significantly lower during holidays and summer holidays.

Investigation of the sociodemographic effects indicates that males have a higher propensity to commute than females. To calculate the overall effect of age and gender, the main effects of age and gender as well as the interaction effects must be tallied. Furthermore, males (25+) make longer commutes than their female counterparts. This observation can be explained by the persistence of traditional patterns: taking care of children still is most frequently done by females, and correspondingly, females better align home and work locations. When employment status is considered, it can be seen that the occupationally active population has a higher likelihood of commuting and spends more time on commuting, than do occupationally inactive people. The higher the position held within a company, the more daily time a person spends on commuting and the higher the probability of commuting. Consequently, executives spend the most time on commuting.

Conclusions that can be drawn from exploring the parameter estimates are that frequent users of public transport (bus, train) commute up to 1.7 times longer than do people who seldom or never use public transport. Daily users of a motorcycle spent on average 35.7% less time on commuting than nondaily users. Also, there is a significant interdependency of travel time expenditure on the remainder of the travel time budget: people making other kinds of trips commute on average 25.4% less than do people who make only commuting trips, and moreover the likelihood of commuting decreases when other type of trips are made. This is a consequence of the substitution effect caused by the travel time frontier, the intrinsic maximum amount of time that people are willing to allocate to travel (*22, 23*).

Time Spent on Shopping Trips

The parameter estimates of the zero-inflated Poisson regression model for predicting travel time expenditure on shopping trips are displayed in Table 5. Recall the distinction between the parameters in the model predicting the mean response λ and the parameters for estimating the probability of the zero-inflation ω. For the analysis, no distinction was made between daily and nondaily shopping, as only 1-day trip-diary data were available. Analysis of temporal effects yields the conclusion that, in general, time spent on shopping trips is less during holidays than during regular days. Saturday appears to be the most preferred day for shopping trips: both the likelihood for performing shopping trips and the travel time expenditure exceed those of other days. An explanation is that on Saturdays there are fewer work-related obligations and more time is available to perform non-work-related activities. The importance of incorporating temporal effects to account for differences in travel time variability was acknowledged by Srinivasan and Guo (*24*) and Habib and Miller (*25*).

Exploration of sociodemographic effects reveals that females have a much larger propensity to perform shopping trips than do males (odds ratio equals 1.37), probably because household-related activities are performed primarily by females (*26*). Assessment of the effect of age is not as straightforward. Adults in the age category 25 to 64 years have the largest probability of performing shopping trips. When the effect of employment status is evaluated, it can be seen that the finding of Gould and Golob (*12*), which indicated that the occupationally active population spends less travel time on shopping than do occupationally inactive people, is more variegated in this study: on the one hand, occupationally active people have a decreased likelihood of performing shopping trips; on the other hand, when they do make the trip, they spent more time than do occupationally inactive people. Although the overall effect remains inconclusive, an important finding is that people performing a liberal profession have a lower likelihood to perform shopping trips (irrespective of self-employed people) and a clearly lower travel time (28% less than executives) than other occupationally active people.

People living in nontraditional situations spend considerably less time on shopping trips. An explanation is that shopping trips for people in living conditions such as rest homes and institutions are performed by staff, instead of by individuals themselves.

One can infer that the degree of urbanization has a decreasing effect on travel time expenditure for shopping trips. A possible reason is the increased number of shopping locations in a more urban context. Furthermore, one could ascertain the interdependence of shopping trips and other kinds of trips. This is again a consequence of the travel time frontier. Note that the interdependency of shopping trips and work trips was incorporated by Lee and Timmermans (*27*).

Time Spent on Leisure Trips

The parameter estimates of the zero-inflated Poisson regression model for predicting travel time expenditure for leisure trips are shown in Table 6. Examination of the temporal effects indicates that both the travel time expenditure on leisure trips and the odds of making these trips are higher during holiday periods and weekends. This can again be explained by the traditional organization of modern society: during weekends and holidays, more time is available to perform leisure activities.

Investigation of sociodemographic effects reveals that males have a higher propensity to perform leisure trips and in general spend more time on leisure trips than do females, which was also demonstrated by Schlich et al. (*26*). People 65 and older are least likely to execute leisure trips and also spend the least time on leisure trips. This is in part because people 65 and older are more likely to have physical disabilities, which limit leisure activities. People who live together have a clearly lower probability and lower travel time expenditure on leisure than people who live alone. Coupling constraints clearly

TABLE 5 ZIP Regression Parameter Estimates for Travel Time Expenditure on Shopping Trips

Parameter	Est.	SE
Poisson Model λ		
Intercept	3.579	0.030
Holiday		
Regular day	0.000	
Public holiday	−0.034	0.011
Summer holiday	−0.001	0.017
Day of week		
Monday	0.000	
Tuesday	0.015	0.021
Wednesday	0.019	0.019
Thursday	0.057	0.019
Friday	−0.013	0.019
Saturday	0.119	0.017
Sunday	−0.021	0.021
Gender		
Male	0.095	0.023
Female	0.000	
Age, years		
6–12	0.300	0.046
13–15	0.394	0.054
16–24	0.197	0.029
25–34	−0.161	0.022
35–44	0.000	
45–54	0.086	0.021
Age, years		
55–64	0.087	0.025
65+	0.064	0.029
Gender & age, years		
Male, 6–12	0.081	0.046
Male, 13–15	−0.736	0.075
Male, 16–24	−0.425	0.041
Male, 25–34	−0.003	0.032
Male, 35–44	0.000	
Male, 45–54	−0.085	0.033
Male, 55–64	−0.031	0.034
Male, 65+	−0.140	0.033
Employment status		
Housekeeping	−0.062	0.023
Unemployed	0.005	0.025
Retired	−0.159	0.025
Disabled	−0.356	0.041
Pupil, student	−0.380	0.033
Worker	0.000	
Employee	−0.009	0.017
Executive	−0.027	0.022
Liberal profession	−0.359	0.061
Self-employed	0.016	0.030
Living conditions		
Alone	0.000	
Others (no partner)	−0.045	0.021
Partner	−0.039	0.017
Partner and others	−0.189	0.017
Other conditions	−0.509	0.075
Degree of urbanization		
Metropolitan area	−0.186	0.022
Urban area	−0.159	0.011
Suburban area	−0.016	0.019
Rural area	0.000	
Use of buses or trams		
Frequently	0.406	0.017
Occasionally	0.087	0.011
Never	0.000	
Use of trains		
Frequently	0.134	0.024
Occasionally	0.053	0.011
Never	0.000	
Other type trips made		
Yes	−0.299	0.010
No	0.000	
Zero Inflation ω		
Intercept	1.345	0.168
Day of week		
Monday	0.000	
Tuesday	−0.043	0.130
Wednesday	−0.295	0.121
Thursday	−0.315	0.123
Friday	−0.423	0.121
Saturday	−1.070	0.116
Sunday	−0.041	0.129
Gender		
Male	0.318	0.069
Female	0.000	
Age, years		
6–12	0.300	0.243
13–15	0.445	0.285
16–24	0.300	0.171
25–34	−0.121	0.105
35–44	0.000	
45–54	0.047	0.110
55–64	−0.008	0.142
65+	0.431	0.183
Employment status		
Housekeeping	−0.870	0.161
Unemployed	−0.840	0.192
Employment status		
Retired	−0.539	0.176
Disabled	−0.239	0.250
Pupil, student	0.000	0.210
Worker	0.000	
Employee	−0.307	0.108
Executive	−0.124	0.145
Liberal profession	0.188	0.337
Self-employed	0.417	0.184
Driving license		
Yes	−0.600	0.105
No	0.000	
Time spent on other trips	0.004	0.001

play an important role here. The importance of incorporating land use and density variables, denoted by Bhat and Gossen (*28*), is also shown by this study: in metropolitan and urban areas, significantly more time is spent on leisure trips, compared to rural areas. Finally, the interdependency of travel time expenditure on differently motivated trips can be observed for leisure trips.

CONCLUSIONS AND FURTHER RESEARCH

This paper showed that sociodemographics, temporal effects, and transportation preferences contribute significantly to understanding variability in daily travel time expenditure. It was shown that the effect of public holidays on daily travel behavior cannot be ignored. The zero-inflated Poisson regression models, which were used to accommodate the Poisson models to the excess of zeros caused by nontravelers, yielded findings that were harmonious with international literature.

The findings reported in this paper should be translated into transportation models. Incorporation of the effect of public holidays on travel demand models will likely result in more-precise travel demand forecasts, and consequently policy makers can develop and fine tune their policy measures on the basis of more-precise assumptions.

Further research should assess the need for accommodating overdispersion in zero-inflated models. The zero-inflated negative binomial approach is a possible framework for tackling both overdispersion and the excess of zeros. A comparison of zero-inflated Poisson regression models with zero-inflated negative binomial regression models would provide a thorough assessment. It also would be worthwhile to compare the suggested modeling approach with the classical techniques, such as Tobit models and hazard-based duration models. Inclusion of social-interaction variables and spatial variables in the analyses could further understanding of differences in travel time expenditure. Moreover, the use of multiday data can improve the analysis by, for instance, differentiating random and routine behavior (*29*). Triangulation of both quantitative (e.g., statistical analysis) and qualitative techniques (e.g., mental models) is a solid approach for gaining insight into the underpinnings of travel behavior.

TABLE 6 ZIP Regression Parameter Estimates for Travel Time Expenditure on Leisure Trips

Poisson Model λ

Parameter	Est.	SE
Intercept	4.575	0.032
Holiday		
Regular day	0.000	
Public holiday	0.361	0.009
Summer holiday	0.415	0.014
Day of week		
Monday	0.000	
Tuesday	0.166	0.020
Wednesday	−0.197	0.021
Thursday	−0.104	0.019
Friday	0.122	0.017
Saturday	0.021	0.016
Sunday	0.056	0.016
Gender		
Male	0.001	0.021
Female	0.000	
Age, years		
6–12	0.166	0.035
13–15	0.192	0.042
16–24	−0.058	0.032
25–34	0.141	0.022
35–44	0.000	
45–54	−0.032	0.023
55–64	−0.038	0.028
65+	−0.544	0.033

Parameter	Est.	SE
Gender & age, years		
Male, 6–12	−0.225	0.031
Male, 13–15	−0.144	0.044
Male, 16–24	0.422	0.031
Male, 25–34	−0.316	0.030
Male, 35–44	0.000	
Male, 45–54	0.133	0.031
Male, 55–64	0.434	0.032
Male, 65+	0.370	0.035
Employment status		
Housekeeping	0.131	0.026
Unemployed	−0.607	0.036
Retired	0.127	0.026
Disabled	0.435	0.038
Pupil, student	−0.006	0.028
Worker	0.000	
Employee	0.311	0.017
Executive	0.237	0.021
Liberal profession	0.556	0.035
Self-employed	0.518	0.024
Living conditions		
Alone	0.000	
Others (no partner)	−0.581	0.020
Partner	−0.154	0.015
Partner and others	−0.385	0.015

Parameter	Est.	SE
Living conditions		
Other conditions	0.970	0.058
Degree of urbanization		
Metropolitan area	0.507	0.017
Urban area	0.100	0.010
Suburban area	0.189	0.015
Rural area	0.000	
Use of buses or trams		
Frequently	−0.162	0.017
Occasionally	−0.110	0.010
Never	0.000	
Use of trains		
Frequently	−0.065	0.020
Occasionally	0.046	0.010
Never	0.000	
Daily use of motorcycle		
Yes	−0.981	0.129
No	0.000	
Driving license		
Yes	−0.292	0.017
No	0.000	
Other type trips made		
Yes	−0.783	0.009
No	0.000	

Zero Inflation ω

Parameter	Est.	SE
Intercept	1.953	0.225
Holiday		
Regular day	0.000	
Public holiday	−0.111	0.081
Summer holiday	−0.264	0.126
Day of week		
Monday	0.000	
Tuesday	0.189	0.155
Wednesday	0.026	0.144
Thursday	−0.190	0.141
Friday	−0.544	0.135
Saturday	−0.868	0.128
Sunday	−1.162	0.132
Gender		
Male	−0.408	0.075
Female	0.000	

Parameter	Est.	SE
Age, years		
6–12	−0.274	0.243
13–15	−0.019	0.272
16–24	−0.312	0.193
25–34	0.017	0.124
35–44	0.000	
45–54	0.194	0.128
55–64	0.258	0.176
65+	0.627	0.232
Employment status		
Housekeeping	−0.383	0.197
Unemployed	−0.350	0.229
Retired	−0.211	0.215
Disabled	0.463	0.353
Pupil, student	−0.558	0.210
Worker	0.000	

Parameter	Est.	SE
Employment status		
Employee	−0.379	0.124
Executive	−0.442	0.159
Liberal profession	−0.688	0.337
Self-employed	0.042	0.203
Living conditions		
Alone	0.000	
Others (no partner)	0.147	0.159
Partner	0.209	0.134
Partner and others	0.368	0.135
Other conditions	2.036	0.706
Driving license		
Yes	−0.364	0.125
No	0.000	
Time spent on other trips	0.005	0.001

REFERENCES

1. Liu, Z., and S. Sharma. Statistical Investigations of Statutory Holiday Effects on Traffic Volumes. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1945,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 40–48.
2. Cools, M., E. A. Moons, and G. Wets. Investigating Effect of Holidays on Daily Traffic Counts: Time Series Approach. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2019,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 22–31.
3. Cools, M., E. Moons, and G. Wets. Investigating the Variability in Daily Traffic Counts Through Use of ARIMAX and SARIMAX Models: Assessing the Effect of Holidays on Two Site Locations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2136,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 57–66.
4. Liu, Z., and S. Sharma. Predicting Directional Design Hourly Volume from Statutory Holiday Traffic. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1968,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 30–39.
5. Liu, Z., and S. Sharma. Nonparametric Method to Examine Changes in Traffic Volume Pattern During Holiday Periods. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2049,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 45–53.
6. Keay, K., and I. Simmonds. The Association of Rainfall and Other Weather Variables with Road Traffic Volume in Melbourne, Australia. *Accident Analysis and Prevention,* Vol. 37, No. 1, 2005, pp. 109–124.
7. Van den Bossche, F., G. Wets, and T. Brijs. Predicting Road Crashes Using Calendar Data. Presented at 85th Annual Meeting of the Transportation Research Board, Washington, D.C., 2006.
8. Zwerts, E., and E. Nuyts. *Travel Behavior Research Flanders (January 2000–January 2001), Part 3A: Analysis of Person Questionnaire* (in Dutch). Departement Architectuur, Provinciale Hogeschool Limburg, Diepenbeek, Belgium, 2004.
9. Agarwal, A. *A Comparison of Weekend and Weekday Travel Behaviour Characteristics in Urban Areas.* MS thesis. University of South Florida, Tampa, 2004.

10. Sall, E. A., and C. R. Bhat. An Analysis of Weekend Work Activity Patterns in the San Francisco Bay Area. *Transportation,* Vol. 34, No. 2, 2007, pp. 161–175.
11. Schwanen, T. The Determinants of Shopping Duration on Workdays in the Netherlands. *Journal of Transport Geography,* Vol. 12, No. 1, 2004, pp. 35–48.
12. Gould, J., and T. F. Golob. Shopping Without Travel or Travel Without Shopping? An Investigation of Electronic Home Shopping. *Transport Reviews,* Vol. 17, No. 4, 1997, pp. 355–376.
13. Frusti, T., C. R. Bhat, and K. W. Axhausen. Exploratory Analysis of Fixed Commitments in Individual Activity-Travel Patterns. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1807,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 101–108.
14. Lanzendorf, M. Mobility Styles and Travel Behavior: Application of a Lifestyle Approach to Leisure Travel. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1807,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 163–173.
15. Cools, M., E. Moons, and G. Wets. Travel Time Expenditure in Flanders: Towards a Better Understanding of Travel Behavior. In *Proc., 10th International Conference on Computers in Urban Planning and Urban Management (CUPUM),* CUPUM, Iguassu Falls, Brazil, 2007.
16. Pendyala, R. M. Time Use and Travel Behavior in Space and Time. In *Transportation Systems Planning: Methods and Applications* (K. G. Goulias, ed.), CRC Press, Boca Raton, Fla., 2003.
17. Lambert, D. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics,* Vol. 34, No. 1, 1992, pp. 1–14.
18. *SAS/STAT 9.1 User's Guide.* SAS Institute, Cary, N.C., 2004.
19. van den Broek, J. A Score Test for Zero Inflation in a Poisson Distribution. *Biometrics,* Vol. 51, No. 2, 1995, pp. 738–743.
20. Simonoff, J. S. *Analyzing Categorical Data.* Springer, New York, 2003.
21. Bhat, C. R., and R. Misra. Discretionary Activity Time Allocation of Individuals Between In-Home and Out-of-Home and Between Weekdays and Weekends, *Transportation,* Vol. 26, No. 2, 1999, pp. 193–209.
22. Banerjee, A., X. Ye, and R. M. Pendyala. Understanding Travel Time Expenditures Around the World: Exploring the Notion of a Travel Time Frontier. *Transportation,* Vol. 34, No. 1, 2007, pp. 51–65.
23. van Wee, B., P. Rietveld, and H. Meurs. Is Average Daily Travel Time Expenditure Constant? In Search of Explanations for an Increase in Average Travel Time. *Journal of Transport Geography,* Vol. 14, No. 2, 2006, pp. 109–122.
24. Srinivasan, K. K., and Z. Guo. Analysis of Trip and Stop Duration for Shopping Activities: Simultaneous Hazard Duration Model System. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1854,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 1–11.
25. Habib, K. M. N., and E. J. Miller. Modeling Individuals' Frequency and Time Allocation Behavior for Shopping Activities Considering Household-Level Random Effects. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1985,* Transportation Research Board of the National Academies, Washington, D.C., 2006, pp. 78–87.
26. Schlich, R., S. Schonfelder, S. Hanson, and K. W. Axhausen. Structures of Leisure Travel: Temporal and Spatial Variability. *Transport Reviews,* Vol. 24, No. 2, pp. 219–237.
27. Lee, B., and H. J. P. Timmermans. A Latent Class Accelerated Hazard Model of Activity Episode Durations. *Transportation Research Part B,* Vol. 41, No. 4, 2007, pp. 426–447.
28. Bhat, C. R., and R. Gossen. A Mixed Multinomial Logit Model Analysis of Weekend Recreational Episode Type Choice. *Transportation Research Part B,* Vol. 38, No. 9, 2004, pp. 767–787.
29. Bhat, C. R., T. Frusti, H. Zhao, S. Schönfelder, and K. W. Axhausen. Intershopping Duration: An Analysis Using Multiweek Data. *Transportation Research Part B,* Vol. 38, No. 1, 2004, pp. 39–60.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Structural Equation Model to Analyze Sociodemographics, Activity Participation, and Trip Chaining Between Household Heads

Survey of Shangyu, China

Min Yang, Wei Wang, Gang Ren, Rui Fan, Bing Qi, and Xuewu Chen

Research in developed countries has demonstrated a significant relationship between activity participation and travel behavior. Less work has been done in China, a typical developing country. Data from an activity–travel survey of 1,205 households in Shangyu, China, were used for a simultaneous structural equation model analysis of sociodemographics, activity participation, and trip chaining between household heads. Model estimation results show that activity participation is strongly correlated to trip chaining characteristics. Additionally, interactions exist between male and female household heads. It was found that the sociodemographics of male and female heads not only directly affected trip chaining but also indirectly influenced trip chaining behavior through activity participation. The findings of the study provide useful insight into the within-personal and cross-personal interactions of activity–travel behaviors between male and female heads in China.

With rapid development in an economy, urban travel demand increases and travel behaviors become more complicated. Facing ever-increasing traffic congestion, people tend to be more concerned with travel demand management (TDM) policies, such as public transit priority and alternative work schedules, than with construction of infrastructure alone (*1, 2*). The making of TDM policies greatly relies on understanding the behavior principles of people's activities and travels (*3–5*). That "travel is a derived demand from activity participation" has long been recognized in studies of travel behavior in developed countries (*6, 7*). Lu and Pas found that complex relationships among sociodemographics, activity participation, and travel behavior exist and that travel behavior could be better explained by including activity participation endogenously in the model, rather than through sociodemographics alone (*8*). Chung and Ahn presented a series of structural equation models, which capture relationships among sociodemographics, activity participation (i.e., time use), and travel behavior for each day during a week in Korea, a developing country (*9*). The need to incorporate the concept of trip chaining of multiple trips into analysis of travel behavior has been widely highlighted (*10, 11*). Kuppam and Pendyala estimated three types of structural equation model systems: one that models relationships between travel and activity participation, another that captures trade-offs between in-home and out-of-home activity durations, and a third that models the generation of complex work trip chains (*12*).

In the mentioned studies, an individual usually was regarded as the primary unit. However, since household members share life resources, many activity–travel decisions of the individual are based on the whole family (*13, 14*). Therefore, models based on individual behavior are inadequate to capture intrahousehold interactions, and hence it is important to incorporate trade-offs of household members. Golob and McNally captured links between activity participation and associated derived travel, links between activities performed by male and female heads, links between types of travel, and time-budget feedback from travel to activity participation. They found that men's work activity participation had a positive effect on women's participation in maintenance activity, an interpersonal and interactivity interaction. In fact, this interaction was the only one that the authors assumed in their model specification (*15*). Golob revealed how the generation of simple and complex trip chains was interrelated with demand for out-of-home and in-home activities. Golob also considered how travel time-budget effects could affect activity demand and trip generation based on household level (*16*). However, his research did not consider interactions between household members. Regarding trip chaining, the research included the numbers of work-related and non-work-related trip chains of the whole household. Cao and Chai applied structural equation models to investigate time allocation of male and female household heads on weekdays and the weekend by using data collected in 1998 from 261 households in Shenzhen, China. The results showed clear household activity roles of Chinese residents: men were dominant in out-of-home activities, and women dominated in-home activities (*17*). Zhang and Chai found significance of activity–travel interaction between household heads positively existed in households in Tianjin, China (*18*). In particular, males dominated outside work-related activities and females dominated outside household-related activities, and they jointly participated in nonwork activities. These two studies concerned only activity participation of household heads without considering trip chaining characteristics. Lee et al. applied simultaneous doubly censored Tobit models to model time-use behavior within the context of household activity participation, to examine time-allocation patterns within household-level trip chaining. Factors found to be

M. Yang, W. Wang, G. Ren, R. Fan, and X. Chen, School of Transportation, Southeast University, No. 2, Sipailou, 210096, Nanjing, China. B. Qi, Department of Civil Engineering, University of Maryland, College Park, 1173 Glenn L. Martin Hall, College Park, MD 20742-3021. Corresponding author: M. Yang, yangmin@seu.edu.cn.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 38–45.
DOI: 10.3141/2157-05

associated with trip chaining behavior included intrahousehold interactions with household types, household structure, and household head attributes (*19*). Yet, trip chaining was only characterized by total travel distance and speed.

Previous studies confirmed that activity participation has a significant relationship with travel behavior, particularly in time use. However, less work has been done to explore the impact of activity participation on trip chaining regarding interactions between household members. Previous studies focused on households in developed countries. There is a great need to formulate relationships of sociodemographics, activity participation, and trip chaining between household heads in developing countries, such as China.

The objectives for this paper are to first capture the links of activity participation and trip chaining between male and female household heads in a Chinese city, and then to confirm the hypothesis that sociodemographics of male and female heads not only directly affect trip chaining but also indirectly influence trip chaining behavior through activity participation. The findings are expected to add to the body of knowledge on relationship of activity participation and travel behaviors.

METHODOLOGY

A structural equation model was applied to estimate a simultaneous model of the interrelationships among sociodemographics, activity participation, and trip chaining behavior, revealing complex intrahousehold interactions between male and female household heads. Of further interest here are the direct and indirect effects of one variable on another, which is a powerful aspect of the structural equation model.

A structural equation model without latent variables has the form (*20–22*)

$$y = By + \Gamma x + \varsigma$$

where

- B = matrix ($p \times p$) of direct effects between pairs of the p endogenous variables,
- Γ = matrix ($p \times q$) of regression effects of the q exogenous variables,
- y = column vector of p endogenous variables,
- x = column vector of q exogenous variables, and
- ς = column vector of the error terms.

Further, denote by Φ the covariance matrix of x and by Ψ the covariance matrix of ς.

Structural equation systems are estimated by covariance-based structural analysis, also called method of moments, in which the difference between the sample covariance and the model-implied covariance matrices is minimized. The fundamental hypothesis for the covariance-based estimation procedures is that the covariance matrix of the observed variables is a function of a set of parameters as shown in the equation: $\Sigma = \Sigma(\theta)$, where Σ is the population covariance matrix of observed variables, θ is a vector that contains the model parameters, and $\Sigma(\theta)$ is the covariance matrix written as a function of θ.

The relation of Σ to $\Sigma(\theta)$ is basic to an understanding of identification, estimation, and assessments of model fit. The matrix $\Sigma(\theta)$ has three components: the covariance matrix of y, the covariance matrix of x with y, and the covariance matrix of x. Then, it can be shown that

$$\Sigma(\theta) = \begin{bmatrix} \Sigma_{yy}(\theta) & \Sigma_{yx}(\theta) \\ \Sigma_{xy}(\theta) & \Sigma_{xx}(\theta) \end{bmatrix} = \begin{bmatrix} (I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)(I-B)^{-1'} & (I-B)^{-1}\Gamma\Phi \\ \Phi\Gamma'(I-B)^{-1'} & \Phi \end{bmatrix}$$

The unknown parameters in B, Γ, Φ, and Ψ are estimated so that the implied covariance matrix $\hat{\Sigma}$ is as close as possible to the sample covariance matrix S. To achieve this, a fitting function $F(S, \Sigma(\theta))$ which is to be minimized, is defined. The fitting function has the properties of being a scalar, greater than or equal to zero, equal to zero if and only if $\Sigma(\theta) = S$, and continuous in S and $\Sigma(\theta)$. Many estimation methods are available. In this research, the maximum likelihood estimation approach is used primarily.

DATA DESCRIPTION

The city selected for this study is Shangyu, a small city located in the middle of the Yangtze River delta region, which is the most active area for economy and culture in China. The city has a population of 204,900, and the area is 111 km^2 with a central district of 18.2 km^2. The government made a travel diary survey of about 4,101 individuals, older than 6 years of age, in 1,564 households, by using random sampling and face-to-face interviews on Thursday, April 27, 2006. The survey had three distinct sections: household characteristics, individual sociodemographics, and travel-activity attributes (*5*). After eliminating missing data and performing logical checking, the authors selected 1,205 households that contained both male and female heads as eligible samples. The exogenous variables include household and individual sociodemographics, and the descriptors of activity participation and trip chaining characteristics are endogenous variables of the model.

There were eight exogenous variables: number of workers, annual household income, and the employment status, age, and educational levels of the male and female heads. Statistical characteristics of exogenous and sociodemographic variables are shown in Tables 1 and 2.

TABLE 1 Statistical Household Characteristics of Exogenous Variables

Variable	Number of Cases	Percentage
No. of workers		
0	126	10.5
1	146	12.0
2	814	67.6
≥3	119	9.9
Annual household income		
<¥10,000	70	5.8
¥10,000–¥20,000	178	14.8
¥20,000–¥50,000	541	44.9
¥50,000–¥100,000	313	26.0
>¥100,000	103	8.5

NOTE: ¥ = The unit of Chinese currency. The currency of China is called the renminbi (RMB). (¥1 = US$0.147 as of July 2010.)

TABLE 2 Statistical Characteristics of Individual Sociodemographic Variables

Variable	Male Head: No. of Cases	Male Head: Percentage	Female Head: No. of Cases	Female Head: Percentage
Employment status				
Worker	935	77.6	726	60.2
Nonworker	270	22.4	479	39.8
Age, years				
20~30	69	5.7	92	7.6
30~40	407	33.8	437	36.3
40~50	377	31.3	358	29.7
50~60	189	15.7	181	15.0
>60	163	13.5	137	11.4
Educational level				
Middle school	307	25.5	502	41.7
High school	491	40.7	489	40.6
Undergraduate	407	33.8	214	17.7

It was found that 67.6% of households have two workers. This statistic can be attributed to the prevailing structure in China of a core household of one couple and one child, where the couple supports the family together. Annual household income is mainly in the middle-to-high groups of 20,000~50,000 and 50,000~100,000 Renminbi (RMB) (1 RMB = US$0.146), which respectively account for 44.9% and 26.0%; this may be because Shangyu is located in the comparatively prosperous area of eastern China. Table 1 shows that the proportion of working males is 16.4% higher than that of females, which indicates males take more social responsibilities. Male and female heads are largely in the 30-to-40 and 40-to-50 age group, respectively 36.3% and 29.7%. There is little gender-based difference of age distribution. Education level is higher for males than for females; the proportion of males with an undergraduate education is 16.1% higher than that of females. This is because for more than 2,000 years, China had been a feudal society in which females were considered inferior to males, and the idea that boys are more important than girls was very common. Although the status of the female has undergone great change and has risen, feudal thought remains. In modern China, the education level of the female is still lower than that of their male counterparts, but the gap is decreasing gradually.

Table 3 shows the 14 endogenous variables, consisting of descriptors of activity participation and trip chaining characteristics for both male and female heads. The travel survey of Shangyu divided the trip purpose into nine categories: work, school, bureaucracy, shopping, social–recreation, serving passengers, personal business, returning home, and returning to work. According to Pas (*23*) and Bowman and Ben-Akiva (*3*), out-of-home activities are divided into three categories: subsistence, maintenance, and leisure. The following classification was applied in Shangyu: subsistence activities (work, school, bureaucracy, and returning to work), maintenance activities (shopping, serving passengers, and personal business), and leisure activities (social–recreation). In-home activity is also considered; however, this activity is not divided into different types because of the limitations of the data. A trip chain is defined as the travel from home to one or more activity locations and back home again. The trip chaining characteristics are defined by descriptors of

TABLE 3 Empirical Analysis of Endogenous Variables

Variable	Average for Working Cases	Average for Nonworking Cases
Activity Participation and Trip Chaining Characteristics of Male Heads		
Number	934	271
In-home	14 h, 37 min	18 h, 35 min
Subsistence	7 h, 54 min	2 h, 51 min
Maintenance	23 min	43 min
Leisure	6 min	47 min
Length (no. of trips)	3.38	3.50
No. of chains	1.54	1.64
Total travel time	60 min	63 min
Activity Participation and Trip Chaining Characteristics of Female Heads		
Number	725	480
In-home	15 h, 3 min	19 h, 59 min
Subsistence	7 h, 35 min	1 h, 22 min
Maintenance	21 min	66 min
Leisure	4 min	33 min
Length (no. of trips)	3.53	3.42
No. of chains	1.65	1.62
Total travel time	57 min	59 min

length (the total trips contained in all trip chains per day), number of chains, and total travel time of all chains for each individual.

It is seen that participation in subsistence activities is distinctly greater than participation in maintenance and leisure for both male and female heads. For example, in working observations, the average subsistence activity of the male is 7 h 54 min, and it is 7 h 35 min for the female. Maintenance and leisure activities are much less, perhaps because a majority of working people in Shangyu spend most of their outside time on work activities during the weekday.

In-home activity participation takes a dominant place, which may be because of the traditional Chinese habit of doing maintenance and leisure activities at home rather than outside of the home. The subsistence, maintenance, and leisure activities of males are all greater than those of females, whereas the in-home activity duration is longer for females than for males. This is consistent with Cao and Chai's finding that "men were dominant in out-of-home activities, but women dominated in-home activities in Shenzhen, China" (*17*).

Table 3 also shows that in working cases, the average trip chaining length for males is 3.38, that is, males take 3.38 trips in a whole day on average. An average of 1.54 chains was found for males, with a total travel time of 60 min. Females' trip chaining length and number of chains are not significantly different from those of males.

MODEL SPECIFICATION AND ESTIMATION RESULTS

Model Specification

The interrelationships among sociodemographics, activity participation, and trip chaining between household heads are examined simultaneously by use of the structural equation modeling methodology. The general model structure is shown in Figure 1. Figure 2

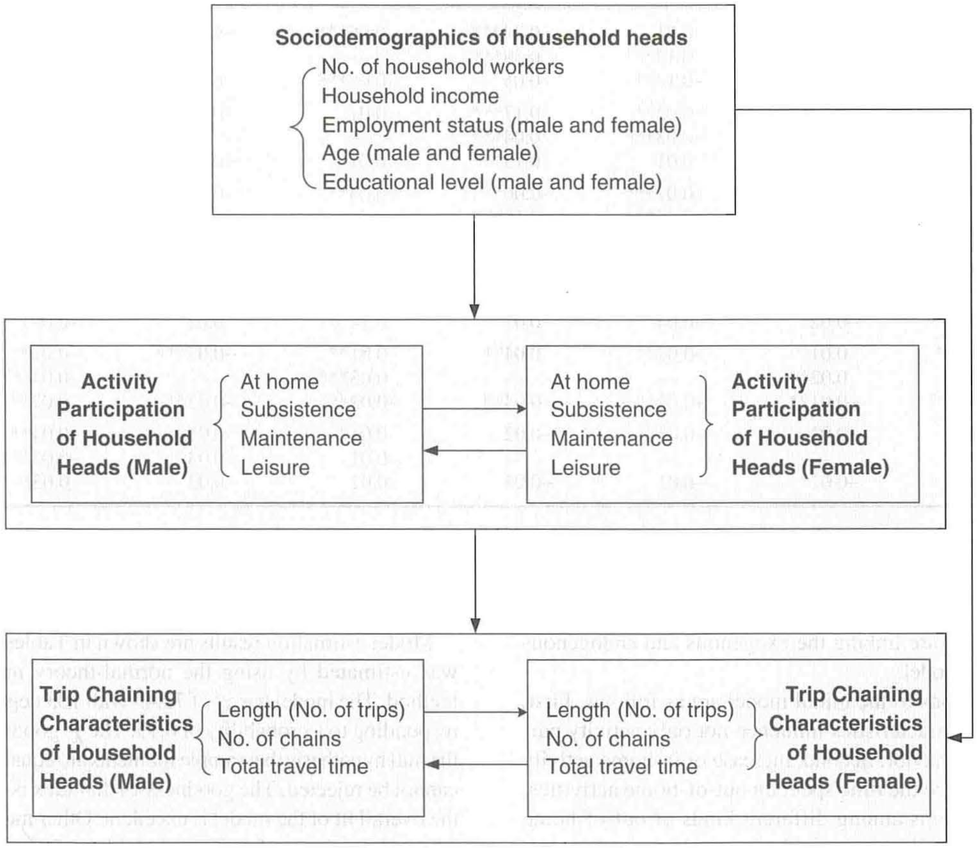


FIGURE 1 Flow diagram of general model structure.

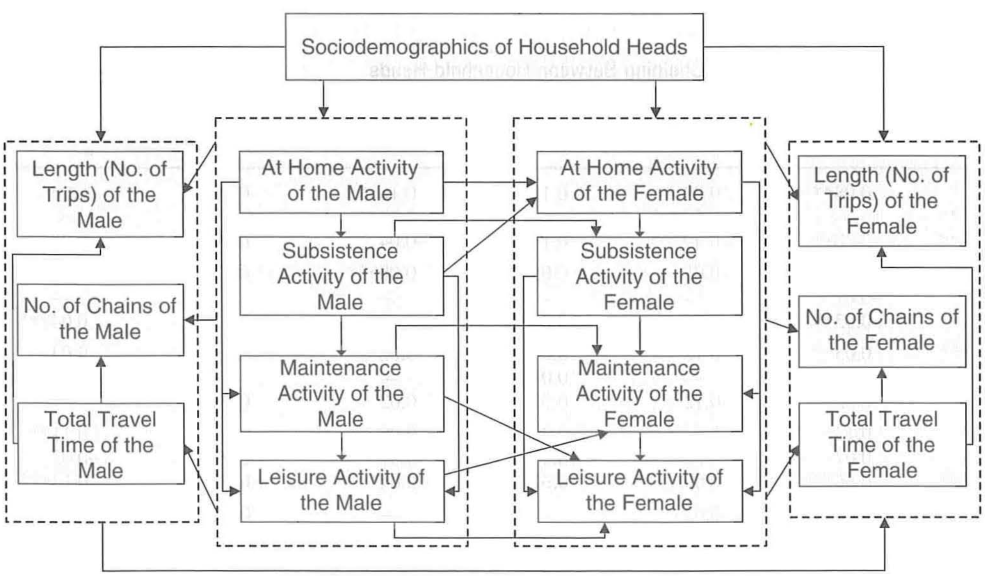


FIGURE 2 Causal structure linking exogenous and endogenous variables in base model.

TABLE 4 Effects of Activity Participation on Trip Chaining Between Household Heads

	In-Home Act. (male)	Subsistence Act. (male)	Maintenance Act. (male)	Leisure Act. (male)	In-Home Act. (female)	Subsistence Act. (female)	Maintenance Act. (female)	Leisure Act. (female)
Trip chain	−0.01	−0.08*	0.01	0.34***	0.05***	−0.13***	0.06**	0.16***
length	−0.03	−0.1**	0.13***	0.39***	—	—	0.02	—
(male)	0.02	0.02	−0.12**	−0.05	0.05***	−0.13**	0.04	0.16**
No. of	0.01	−0.03**	−0.02**	0.17***	0.01	−0.04**	0.01	0.07***
chains	0.04***	0.03***	−0.03***	0.04***	—	—	—	—
(male)	−0.03***	−0.06***	0.01	0.13***	0.01	−0.04**	0.01	0.07***
Total travel	−0.02	−0.02	−0.07***	−0.06***	0.04***	−0.05***	0.02*	0.01
time	−0.15***	−0.15***	−0.08***	−0.15***	—	—	—	—
(male)	0.13***	0.13***	0.01	0.09***	0.04***	−0.05***	0.02*	0.01
Trip chain	0.07***	−0.02	−0.04	−0.01	0.03*	−0.13***	0.02	0.14
length	—	—	—	—	−0.11***	−0.15***	0.11***	0.14
(female)	0.07***	−0.02	−0.04	−0.01	0.14***	0.02	−0.09***	—
No. of	0.03**	0.01	−0.02**	0.04**	0.01**	−0.03***	−0.02**	0.05
chains	−0.02*	0.02*	—	—	0.03***	—	−0.04***	−0.02
(female)	0.05***	−0.01**	−0.02**	0.04**	−0.02**	−0.03**	0.02**	0.07
Total travel	0.01	−0.02	−0.01	−0.02	0.01*	−0.01	−0.04***	−0.04
time	0.03	—	—	—	−0.01	−0.03	−0.07***	−0.08***
(female)	−0.02	−0.02	−0.01	−0.02	0.02	0.02	0.03*	0.04

NOTE: Total effect (level of significance); direct effect (level of significance); and indirect effect (level of significance).
***$p < .01$; **$p < .05$; *$p < .1$.

shows the causal structure linking the exogenous and endogenous variables in the base model.

The basic assumptions to the initial model are as follows. First, sociodemographics characteristics influence not only activity participation but travel behavior. Second, increase of in-home activity participation will reduce the time spent on out-of-home activities, and there are interactions among different kinds of out-of-home activities. Third, household and individual sociodemographics not only influence trip chaining directly, but have an indirect impact on trip chaining through activity participation. Last, interactions exist between male and female household heads.

Model estimation results are shown in Tables 4 and 5. The model was estimated by using the normal-theory maximum likelihood method. The model has χ^2 of 70.43 with 103 degrees of freedom corresponding to a probability of 0.99. The χ^2 goodness of fit shows that the null hypothesis that sample moments are equal to implied moments cannot be rejected. The goodness-of-fit index is 1.00, indicating that the overall fit of the model is excellent. Other measures of fit, such as adjusted goodness-of-fit index (AGFI = 0.99) and root-mean-square error of approximation (RMSEA = 0.000), are found to be acceptable by model fit criteria for structural equations models. Hoelter's critical *N* statistic is found to be 2669.03 (≥ 200 is acceptable), which is the

TABLE 5 Effects of Sociodemographics on Trip Chaining Between Household Heads

	No. of Workers	Household Income	Employment Status (male)	Age (male)	Educational Level (male)	Employment Status (female)	Age (female)	Educational Level (female)
Trip chain	−0.02	0.08**	0.09	0.13***	0.16***	0.09	0.05	0.04
length	0.19***	—	0.58***	—	0.2***	—	—	—
(male)	−0.21***	0.08**	−0.49***	0.13***	−0.04	0.09	0.05	0.04
No. of	−0.01	0.01	0.05	0.05***	0.08***	0.02	0.04***	0.02
chains	—	−0.02	—	—	—	—	—	—
(male)	−0.01	0.03**	0.05	0.05***	0.08***	0.02	0.04***	0.02
Total travel	−0.09***	0.03**	0.12***	0.06**	0.02	0.02	0.03	−0.04
time	−0.10***	—	—	0.04*	—	—	0.03	−0.05*
(male)	0.01	0.03**	0.12***	0.02	0.02	0.02	—	0.01
Trip chain	−0.09	0.13**	0.1	−0.02	0.05	0.1	0.11**	0.14*
length	0.02	0.07	0.05	−0.03	0.09	0.52***	−0.02	0.21***
(female)	−0.11***	0.06***	0.05	0.01	−0.04*	−0.42***	0.13***	−0.07**
No. of	−0.02	0.04*	0.02	—	—	0.02	0.04	0.06**
chains	—	—	−0.11***	0.02	−0.05***	—	—	—
(female)	−0.02	0.04*	0.13**	−0.02	0.05*	0.02	0.04	0.06**
Total travel	−0.06**	0.02	−0.02	−0.01	0.06**	0.08*	0.07***	0.04
time	−0.03	—	—	−0.03	0.03	—	0.05*	—
(female)	−0.03	0.02	−0.02	0.02	0.03	0.08	0.02	0.04*

NOTE: Total effect (level of significance); direct effect (level of significance); and indirect effect (level of significance).
***$p < .01$; **$p < .05$; *$p < .1$.

sample size at which the value of the fitting function F_{ML} would lead to the rejection of the null hypothesis, H_0 [i.e., $\Sigma = \Sigma(\theta)$], at a chosen significance level.

Effect of Activity Participation on Trip Chaining Between Household Heads

Table 4 shows the total, direct, and indirect effects of activity participation on trip chaining characteristics between male and female household heads.

It can be seen from the total effect that males' trip chaining length (number of total trips) decreases as subsistence activity duration increases, with the value of −0.08. However, longer leisure activity duration increases the trip chaining length with total effect of 0.34. It was also found that number of chains of the male head is negatively related to subsistence and maintenance activity participation (with total effects of −0.03 and −0.02, respectively) but positively related to leisure activity duration (with total effect of 0.17). These findings agree with the research results of developed countries, indicating that subsistence activity has some constraints in travel, whereas leisure activity induces travel. It is possible that subsistence activities have spatial-temporal constraints, and the travel involved in subsistence activities should satisfy these constraints; that is, to some extent, subsistence activities restrict travel. Flexible activities such as maintenance and leisure activities are relatively discretional. The existence of flexible activities allows one freedom in travel generation. Leisure activity affects the length and number of trip chains to a larger degree for males than does subsistence activity. All kinds of activities (in-home, subsistence, maintenance, and leisure) have negative total effects on total travel time of trip chaining, with the values of −0.02, −0.02, −0.07, and −0.06, respectively. That is, an average increase of 1 h in in-home or subsistence activity will correspondingly decrease the total travel time of trip chaining by 1.2 min. Meanwhile, a 1 h increase in maintenance or leisure activity may result in a decrease of 4.2 min or 3.6 min in the total travel time, respectively. This is significantly different from what is found in Western developed countries, where the duration of out-of-home activities is generally positively related to travel time (*8*). The main reason for this is that maintenance and leisure activities are distributed extensively in China, and it takes a very short time to get to the locations for these activities. Thus, it is feasible to put more time into the activities. This brings about the negative relationship between activity participation and total travel time of trip chaining. This finding is consistent with that of Cao and Chai (*17*).

Table 4 also shows that the length and number of trip chains of females are negative to their subsistence and maintenance activities, which shows the same tendency as male heads. However, in-home activity causes a positive effect to trip chaining for females. In a small city like Shangyu, the subtour "returning home at noon" (hwhwh) takes place more frequently for women, and the time they stay at home is relatively longer. With the existence of a subtour, travel time is relatively longer. As a result, the relationship between the time staying at home and travel time is a positive correlation. As for male heads, maintenance activity negatively affects the total travel time of trip chaining of females with a coefficient of −0.04. This suggests that an average increase of 1 h in maintenance activity would reduce the travel time by 2.4 min for females.

In addition, there are interactions between male and female household heads. The males' in-home activity participation causes a total effect of 0.07 to the females' trip chaining length and a total effect of 0.03 to the number of trip chains. It is also seen that males' trip chaining length is positively correlated to activity participation by the female in home. This finding reveals that if one household head spends more time at home, the other would gain more opportunities to go outside. The data suggest that if females increase their in-home activity participation, males' total travel time will also increase. Yet, females' participation in subsistence activity causes negative effects on length and number of trip chains for males. However, the effects of males' activity participation are not noticeable on females' trip chaining. The interactions between male and female heads illustrate the distribution mechanism in sharing the household tasks and the association or substitution effects of their activity–travel patterns.

Influence of Sociodemographics on Trip Chaining of Indirect Effect from Activity Participation

Table 5 indicates the impact of household and individual sociodemographics of male and female heads on trip chaining by containing the indirect effect from activity participation. From total effects, it can be seen that when the number of working members increases, the length, number, and total time of trip chaining will decrease for both male and female heads. However, positive total effects of annual household income on trip chaining characteristics are found for both heads. A possible reason is that household heads with higher incomes are generally more active in outside activities. Additionally, the trip chaining characteristics of females are affected by household income more than are those of males.

There is significant difference between the total effect (−0.02) and the direct effect (0.19) of the variable of number of working members on males' trip chaining length. This clearly comes from the indirect effect with −0.21. Tracing the path diagram of the model shows that the negative indirect impact is mainly from participation in subsistence and leisure activities, seen in Table 6. This reveals that overworking of household members restricts travel, so more working members result in less leisure activity and thus the number of trips is reduced. This accords with the conclusion from analysis of the relationship between activity participation and trip chaining characteristics in Table 4.

Table 5 also shows that working, older, and better-educated male heads have a positive effect on all trip chaining characteristics, including length, number, and total time of trip chains. There is a similar tendency for female heads. This indicates that activity–travel probability increases along with age and education level, which may bring more chances for social activities. A comparison of the direct

TABLE 6 Indirect Effect of Number of Workers on Male Trip–Chain Length

Mediating Variable	Indirect Effect
In-home activity participation (male)	0.0114
Subsistence activity participation (male)	−0.0910
Maintenance activity participation (male)	−0.0236
Leisure activity participation (male)	−0.0702
In-home activity participation (female)	0.0044
Leisure activity participation (female)	−0.0168
Other trip chaining descriptors	−0.0242
Total	−0.21

TABLE 7 Indirect Effect of Employment Status of Males on Trip Chain Length

Mediating Variable	Indirect Effect
In-home activity participation (male)	−0.0615
Subsistence activity participation (male)	−0.3030
Maintenance activity participation (male)	−0.0962
Leisure activity participation (male)	−0.2535
Maintenance activity participation (female)	0.0239
Other trip chaining descriptors	0.2003
Total	−0.49

and total effects shows that the indirect effect coming from activity participation plays an important role. For instance, a positive direct effect of employment status on males' trip chaining length is 0.58; however, the indirect effect is −0.49, leading to a total effect of only 0.09. A tracing of the path diagram reveals that subsistence and leisure activities mainly contribute to this negative indirect effect, as Table 7 shows. This may be because devoting more time to work restricts travel, whereas less participation in leisure activities decreases the number of trips included in the trip chains for males.

In addition, no significant cross-personal interactions exist in the total effects of sociodemographics on trip chaining between male and female heads. However, some interactive indirect effect from activity participation was found. For example, working males have a negative direct effect of −0.11 on the number of chains of females, and the total effect is positive with the value of 0.02. This reflects that working males increase the number of chains for females. This is contributed by the indirect effect of 0.13, which is derived mainly from working males' in-home and subsistence activities as presented in Table 8. A possible explanation is that if males spend more time at home, female heads would have more opportunities for outside activities and travel. Meanwhile, when male heads engage in a longer subsistence activity, females must take responsibility for more out-of-home maintenance activities, and this increases the number of trip chains. This finding is consistent with the work of Golob and McNally (*15*).

CONCLUSIONS AND FUTURE DIRECTIONS

This research establishes a structural equation model to simultaneously explore the relationships among sociodemographics, activity participation, and trip chaining characteristics between household heads in Shangyu, China. Model estimation results indicate that activity participation not only exerts within-personal impact on trip chaining but also causes cross-personal interactions between male and female heads. Further, it confirms that sociodemographics not only affect trip chaining characteristics directly but also have indirect effects through activity participation.

TABLE 8 Indirect Effect of Employment Status of Males on Number of Trip Chains of Females

Mediating Variable	Indirect Effect
In-home activity participation (male)	0.0820
Subsistence activity participation (male)	0.0606
Maintenance activity participation (male)	0.0222
Leisure activity participation (male)	−0.0260
Other trip chaining descriptors	−0.0088
Total	0.13

Participation in subsistence activity restricts both the trip length and the number of trip chains for both male and female heads, whereas leisure activity induces trip chains. However, in-home activity participation positively affects the trip chaining of females. Different from what is shown by research results from developed countries, the four kinds of activity participation all exert negative effects on the total travel time of trip chains for both heads. Cross-personal interactions were found to exist between male and female heads. Males' in-home activity participation causes a positive effect on the length and number of trip chains of females, and the same tendency is shown for the effect of females on males. This finding reveals that if one household head spends more time at home, the other would gain more opportunities to go outside.

Additionally, it can be seen from the total effect that when the number of working members increases, the length, number, and total time of trip chaining will decrease for both male and female heads. However, household income positively affects trip chaining for both heads. It was also found that older working male heads with higher education levels positively affect all females' trip chaining characteristics, and vice versa. Tracing the path diagram of the model clearly shows that the indirect effects of sociodemographics on trip chaining characteristics are mainly derived from activity participation.

This research enriches the study of the relationship among sociodemographics, activity participation, and trip chaining between household members. Furthermore, it provides useful insight into the activity–travel behaviors of male and female household heads in China. The authors of this paper are expected to promote the application of the concept that travel is derived from activity in travel demand management in Chinese cities. The study of relationships among sociodemographics, activity participation, and travel behavior can help planners and engineers to evaluate TDM policies by considering the links between intrahousehold interaction and individual travel. The findings of the study can be used to seek analytical methods and establish analytical mode for TDM.

This research can be advanced in the following aspects. First, in addition to analyzing the effect of activity participation on trip chaining characteristics, researchers could conduct specific studies on trade-off relationships between different kinds of activities. Then the in-home activities could be subdivided on the condition of more detailed data. Second, rather than including only overall trip chaining characteristics, future research may consider different types of trip chaining. Third, because of data limitations, interaction of activity–travel behaviors between children and household heads was not considered; this may be an important aspect for enriching research on household decision making.

ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation and the National Basic Research Program of China. The authors thank Wei Wang of the Massachusetts Institute of Technology and Mark L. Franz of the University of Maryland for their help in revising the English expressions used in this paper.

REFERENCES

1. Juan, Z. C., F. Zong, and X. Y. Fu. Literature Review on TDM Strategies Implementation and Evaluation. *Journal of Chinese Soft Science,* Vol. 1, 2007, pp. 156–160.
2. Sundo, M. B., and S. Fujii. The Effects of a Compressed Working Week on Commuters' Daily Activity Patterns. *Transportation Research Part A,* Vol. 39, No. 10, 2005, pp. 835–848.
3. Bowman, J. L., and M. Ben-Akiva. Activity-Based Disaggregate Travel Demand Model System with Activity Schedules. *Transportation Research Part A,* Vol. 35, No. 1, 2000, pp. 1–28.
4. Bhat, C. R. Modeling Commute Activity-Travel Pattern of Workers: Formulation and Empirical Analysis. Presented at 78th Annual Meeting of the Transportation Research Board, Washington, D.C., 1999.
5. Yang, M., W. Wang, X. Chen, T. Wan, and R. Wu. Empirical Analysis of Commute Trip Chaining: Case Study of Shangyu, China. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2038,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 139–147.
6. Pendyala, R. M., and K. G. Goulias. Time Use and Activity Perspectives in Travel Behavior Research. *Transportation,* Vol. 29, No. 1, 2002, pp. 1–4.
7. Timmermans, H., T. Arentze, and C. H. Joh. Analysing Space-Time Behavior: New Approaches to Old Problems. *Progress in Human Geography,* Vol. 26, No. 2, 2002, pp. 175–190.
8. Lu, X. D., and E. I. Pas. Socio-demographics, Activity Participation and Travel Behavior. *Transportation Research Part A,* Vol. 33, No. 1, 1999, pp. 1–18.
9. Chung, J. H., and Y. S. Ahn. Structural Equation Models of Day-To-Day Activity Participation and Travel Behavior in a Developing Country. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1807,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 109–118.
10. Wallace, B., J. Barnes, and G. S. Rutherford. Evaluating the Effects of Traveler and Trip Characteristics on Trip Chaining, with Implications for Transportation Demand Management Strategies. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1718,* TRB, National Research Council, Washington, D.C., 2000, pp. 97–106.
11. McGuckin, N., J. P. Zmud, and Y. Nakamoto. Trip Chaining Trends in the United States: Understanding Travel Behavior for Policy Making. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1917,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 199–204.
12. Kuppam, A. R., and R. M. Pendyala. A Structural Equations Analysis of Commuters' Activity and Travel Patterns. *Transportation,* Vol. 28, No. 1, 2001, pp. 33–54.
13. Bhat, C. R., and R. M. Pendyala. Modeling Intra-household Interactions and Group Decision-making. *Transportation,* Vol. 32, No. 5, 2005, pp. 443–448.
14. Yang, M., W. Wang, and X. W. Chen. Activity Pattern Choice of Work Commuting Trip by Workers. *Journal of Southwest Jiaotong University,* Vol. 44, No. 2, 2009, pp. 274–279.
15. Golob, T. F., and M. G. McNally. A Model of Activity Participation and Travel Interactions Between Household Heads. *Transportation Research B,* Vol. 31, No. 3, 1997, pp. 177–194.
16. Golob, T. F. A Simultaneous Model of Household Activity Participation and Trip Chain Generation. *Transportation Research B,* Vol. 34, No. 5, 2000, pp. 355–376.
17. Cao, X. Y., and Y. W. Chai. Gender Role-Based Differences in Time Allocation: Case Study of Shenzhen, China. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2014,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 58–66.
18. Zhang, W. J., and Y. W. Chai. Theories and Confirmed Model of Urban Resident's Travel Demand: Considering Intrahousehold Interaction. *Acta Geographica Sinica,* Vol. 63, No. 12, 2008, pp. 1246–1256.
19. Lee, Y., M. Hickman, and S. Washington. Household Type and Structure, Time-use Pattern, and Trip chaining Behavior. *Transportation Research Part A,* Vol. 41, No. 10, 2007, pp. 1004–1020.
20. Fox, J. Effect Analysis in Structural-Equation Models II: Calculation of Specific Indirect Effects. *Sociological Methods and Research,* Vol. 14, No. 1, 1985, pp. 81–95.
21. McDonald, R. P., and M. R. Ho. Principles and Practice in Reporting Structural Equation Analyses. *Psychological Methods,* Vol. 7, No. 1, 2002, pp. 64–82.
22. Tomarken, A. J., and N. G. Waller. Structural Equation Modeling: Strengths, Limitations, and Misconceptions. *Annual Review of Clinical Psychology,* Vol. 1, No. 1, 2005, pp. 31–65.
23. Pas, E. I. Weekly Travel-Activity Behavior. *Transportation,* Vol. 15, No. 1–2, 1988, pp. 89–109.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Scobit-Based Panel Analysis of Multitasking Behavior of Public Transport Users

Junyi Zhang and Harry Timmermans

With a focus on the multitasking behavior of public transportation users during travel, an examination is made of factors affecting activity participation along the axis of travel time. The probability of participation in an activity is represented by using a scobit (or a skewed logit) model, within which the widely used logit model is nested with the help of a skewness parameter. With this skewness parameter, it is not necessary to assume that individuals with a probability of 0.5 for performing an activity are most sensitive to changes in travel time or other influential factors. This analysis is the first attempt to apply the scobit model to transportation issues. An empirical analysis was conducted by using data (523 individuals) collected in Hiroshima City, Japan, in December 2008. Because multitasking behavior along the axis of travel time may be interrelated, the scobit model is extended to simultaneously incorporate the influences of state dependency and the remaining travel time as well as the other influential factors, by dealing with the data as panel data. As a result, a scobit-based panel model was developed. Model estimation results confirm the effectiveness of the scobit model. It was further revealed that introduction of a heterogeneous skewness parameter is more effective for representing activity participation than assumption of a homogeneous skewness parameter. Calculation results of travel time elasticity show that activity participation is sensitive to change in the travel time up to the first half of travel time and becomes less sensitive to travel time afterward.

In recent years, multitasking, which is broadly defined as simultaneous engagement in multiple activities (*1–3*), has been studied from various perspectives. Daily life often requires performing multiple tasks to save time (*4*), and making effective use of time is another reason for multitasking. Interest in multitasking behavior in the field of transportation is based on several motivations (*5*). First, information technology has substantially increased opportunities to work or engage in leisure activities while traveling, blurring the boundaries between work and travel. Second, the multitasking discussion considers whether the utility of travel time is necessarily negative because travelers can conduct activities while traveling. It might be expected that the longer the travel time, the higher the probability that some activities will be conducted during travel; however, this does not necessarily mean that at some very long distance, the utility of travel becomes positive. Third, analysis of multitasking may shed light on economic and social topics, such as gender equity in leisure time, the economic value of work while traveling and work at home, the nature of paid and unpaid work, the fragmentation of time, and the contamination of leisure time. Fourth, because activities conducted while traveling may substitute the same or other activities conducted at other times, multitasking perhaps should be an integral component of activity-based models. Finally, multitasking has some policy implications. If participation in activities while using public transportation systems is a good reason for choosing public transit rather than car travel, or if such participation could reduce resistance to use of public transportation systems, policy makers could improve traveling environments to meet such needs. For example, in some major Japanese cities such as Tokyo, Nagoya, and Osaka, operators of public transportation systems provide news, weather forecasts, and entertainment programs via in-vehicle television. On some platform in major Tokyo stations, for example, one can use wireless LAN systems to access the Internet.

Studies have examined the importance of information and communication technologies (ICTs) for travel and activities. Mokhtarian et al. argued that multitasking ability is one of the major characteristics of ICTs that support their increasing popularity as facilitators and conceptually explored the potential impact of ICTs on leisure activities and associated travel (*6*). The limited human ability to engage in more than one task at a time can be improved through the use of the Internet and mobile phone (*7*). Because of multitasking made possible by wireless ICTs, the analytical distinction between stationary activities, physical movement, and communication increasingly feels like a straitjacket (*8*). In a study of the impact of the Internet, Kenyon and Lyons empirically confirmed that multitasking occurs daily for all individuals (*1*). They argued that failure to consider multitasking leads to underreporting of key activities, and lack of awareness of multitasking could lead to flawed measurement and misrepresentation of behavior change (*1*). Kenyon also showed that focus-group participants described their low awareness of multitasking before using the diary instrument, suggesting that without the diary prompt, they would have substantially underreported their participation in secondary activities (*9*).

Another major research category has explored the possibility that travel time has a positive utility. Jain and Lyons suggested that travel time, at least for individuals, can sometimes be perceived and experienced as a gift rather than a burden (*10*). They identify two forms of travel time experience from which positive utility can be derived, transition time and time out—both facilitated or supported by a third notion: equipped time. Ettema and Verschuren used a stated preference approach to investigate the relationship between multitasking during travel and the valuation of travel time and found that commuters who dislike engaging in activities or who read for their work while commuting have a higher value of travel time, and commuters who listen to music while commuting have a lower value of travel time (*2*).

J. Zhang, Transportation Engineering Laboratory, Graduate School for International Development and Cooperation, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima, 739-8529, Japan. H. Timmermans, Faculty of Architecture, Building, and Planning, Urban Planning Group, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands. Corresponding author: J. Zhang, zjy@hiroshima-u.ac.jp.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 46–53.
DOI: 10.3141/2157-06

Focusing on public transportation, Timmermans and Van der Waerden found that time allocation to various activities during travel depends primarily on sociodemographics such as gender, age, travel party, and race, and they confirmed the effects of contextual variables such as travel duration and time of day, which are particularly relevant for discriminating between different segments of activities while traveling (*3*). Van der Waerden et al. estimated that the most important variables influencing the probability of distinguished tasks are presence of accompanying persons, age of traveler, travel day, type of train, crowdedness of train, and time of day (*11*).

If multitasking while traveling could significantly change the value of travel time, economic transportation schemes should be reassessed and redefined (*3*), and more evidence is needed. Studies on multitasking in transportation are few. This study attempts to provide insight into the analysis of multitasking behavior by focusing on the generation of multitasking behavior during the use of public transport systems. First, how to represent the generation of an activity over the course of travel in a more useful way is explored. Second, since the length of travel time may influence participation in activities, the influences of travel time on activity generation during travel is examined.

Logit models or probit models have often been applied to activity generation. These two models assume that sensitivity to changes in explanatory variables is highest for individuals who are invariant between choosing or not choosing to perform an activity (probability of 0.5). However, in the real world, this assumption may not hold. In other words, the trip makers most sensitive to changes in an explanatory variable may have an initial probability of activity generation. To overcome this potential shortcoming of the logit and probit model, this study applied a scobit (or skewed logit) model (*12*), within which the widely used logit model is nested. Different from the logit model, a skewness parameter is introduced into the scobit model. The skewness parameter is positive, and when it is equal to one the scobit model returns to the logit model. It is expected that the use of the scobit model better will allow one to identify the sensitivity of travelers to changes in influential factors. The multitasking behavior of public transportation users is analyzed by using data (valid sample size: 523 individuals) collected in Hiroshima City, Japan, in December 2008.

MODELING FRAMEWORK

It is expected that exploring whether an activity is generated during travel could contribute to a better understanding of travel time savings. A trip maker might repeatedly perform these two types of activities over the course of the limited travel time. To capture such behavioral changes, panel analysis may be applicable (*13*). This study uses a set of multitasking behavior data from a trip maker as panel data. Unlike standard panel analysis, however, the panel data in this study have different waves (equal to the number of activities performed during travel) across individuals, and these waves are not discrete but are connected. Also, the total time period is much shorter.

Scobit-Based Modeling

First, define the utility u_{nt} that trip maker n performs an activity ($y_{nt} = 1$) or does not perform an activity ($y_{nt} = 0$) at the tth wave (i.e., time period) as follows:

Activity participation:

$$u_{nt}(y_{nt}=1) = v_{nt}(1) + e_{nt}(1) \tag{1}$$

Nonactivity participation:

$$u_{nt}(y_{nt}=0) = v_{nt}(0) + e_{nt}(0) \tag{2}$$

where

$u_{nt}(1)$, $u_{nt}(0)$ = utility functions of activity participation and nonactivity participation,
$v_{nt}(1)$, $v_{nt}(0)$ = deterministic terms of $u_{nt}(1)$, $u_{nt}(0)$, and
$e_{nt}(1)$, $e_{nt}(0)$ = corresponding error terms, respectively.

Then the probability $P_{nt}(y_{nt} = 1)$ that trip maker n performs an activity at the tth wave can be described as

$$\begin{aligned} P_{nt}(y_{nt}=1) &= \Pr\left(u_{nt}(y_{nt}=1) > u_{nt}(y_{nt}=0)\right) \\ &= \Pr\left(e_{nt}(1) - e_{nt}(0) > v_{nt}(0) - v_{nt}(1)\right) \end{aligned} \tag{3}$$

Assume a new error term $\epsilon_{nt} = e_{nt}(1) - e_{nt}(0)$, which follows a distribution with the distribution function $F(\epsilon_{nt})$. Without loss of generality, it is further assumed that $v_{nt}(0) = 0$. Then the probabilities $P_{nt}(y_{nt} = 1)$ and $P_{nt}(y_{nt} = 0)$ can be derived as

$$P_{nt}(y_{nt}=1) = 1 - F\left(-v_{nt}(1)\right) \tag{4}$$

$$P_{nt}(y_{nt}=0) = F\left(-v_{nt}(1)\right) \tag{5}$$

Such a utility-based choice model usually assumes that the deterministic term $v_{nt}(1)$ is a linear function of explanatory variables (x_{ntk} = the kth variable with parameter β_k):

$$v_{nt}(1) = \sum_k \beta_k x_{ntk} \tag{6}$$

Policy makers or analysts always need to know the marginal effect of a change in x_{ntk} on $P_{nt}(y_{nt} = 1)$. Such marginal effect can be derived as follows:

$$\frac{\partial P_{nt}(y_{nt}=1)}{\partial x_{ntk}} = f\left(-\sum_k \beta_k x_{ntk}\right)\beta_k \tag{7}$$

where $f(\cdot)$ is the probability density function of $F(\epsilon_{nt})$.

It is obvious that $\partial P_{nt}(y_{nt} = 1)/\partial x_{ntk}$ will depend not only on β_k but also on the value of x_{ntk}, and in particular $f(\cdot)$. If a normal or Weibull distribution is assumed, then $f(\cdot)$ will have a maximum at $\Sigma_k\beta_k x_{ntk} = 0$. This means that any given variable x_{ntk} will have its greatest effect on those individuals for which $\Sigma_k\beta_k x_{ntk}$ is closest to 0, or for which $\partial P_{nt}(y_{nt} = 1)$ is closest to 0.5. However, if individuals with initial probability other than 0.5 are those most sensitive to the change, then the logit or probit model would result in a misspecification and, consequently, biased inferences about the marginal effects of changes in any explanatory variable. In this sense, it is necessary to adopt a more general distribution, which could allow the highest sensitivity to changes in variables at any initial probability. To meet this

requirement, this study applies the scobit (or skewed logit) model (*12*). This model can be obtained by assuming the following $F(\epsilon_t)$, which is a Burr-10 distribution (*14*):

$$F(\epsilon_t) = \frac{1}{(1+\exp(-\epsilon_t))^{\alpha}} \tag{8}$$

where α is a parameter used to measure the skewness of the Burr-10 distribution.

With the distribution function $F(\epsilon_t)$ defined, the probabilities of activity participation during travel can be derived as

$$P_t(y_t = 1) = 1 - F(-v_t(1)) = 1 - \frac{1}{(1+\exp(v_t))^{\alpha}} \tag{9}$$

$$P_t(y_t = 0) = 1 - P_t(y_t = 1) = \frac{1}{(1+\exp(v_t))^{\alpha}} \tag{10}$$

The Burr-10 distribution satisfies the condition that $f(\cdot)$ does not attain a maximum only when $F(\cdot) = 0.5$, and it is defined for $-\infty < \epsilon_t < \infty$. When α is equal to 1, Equations 9 and 10 return to the logit model. Thus, the popular logit model is nested within the scobit model. The scobit model is also called the skewed logit model, because it allows for a skewed response curve, which is different from the symmetric curve (symmetric about zero) derived from the logit model. Considering the popularity of the logit model, it is worth examining the applicability of the scobit-based model to the analysis of multitasking while traveling.

Specification of Utility Function

Decision-making situations involved in multitasking behavior during travel are illustrated in Figure 1. During travel, a trip maker decides to perform an activity A_{nt} or not at time period (or wave) t_n. Since different trip makers may perform different numbers of activities, the number of time periods and the duration of each period may vary across individuals. Such sequential and repeated choice behavior could be represented within a panel analysis framework.

The deterministic term $v_{nt}(1)$, that is, the nonstochastic utility of participating in an activity, may be influenced by several groups of variables. First, it could be influenced by activity participation in previous time periods. For example, if a trip maker already read a book during a previous period, to relax he or she may wish to listen to music during other periods, or may feel tired and not do anything. Such behavioral dependence could be described by the concept of state dependence. Second, the remaining travel time when one is deciding to perform an activity during travel may influence the utility. This can be used to examine whether longer travel times will increase the probability of multitasking. Third, $v_{nt}(1)$ could be affected by other types of activity-specific variables (e.g., whether to stand or take a seat during the use of a public transportation system, whether to travel alone or with other persons, and time of a day) and activity-generic attributes (e.g., age and sex). Summarizing, $v_{nt}(1)$ can be defined as follows:

$$v_{nt}(1) = \mu + \rho y_{nt-1} + \gamma\tau_{nt} + \sum_k \beta_k x_{ntk} + \sum_s \kappa_s z_{ns} \tag{11}$$

where

μ = constant term,
y_{nt-1} = dummy variable (equal to 1 when an activity is performed at wave $t-1$, otherwise 0),
ρ = parameter indicating state dependence,
τ_{nt} = remaining travel time when deciding to perform an activity during travel,
γ = parameter of τ_{nt},
x_{ntk} = kth activity-specific attribute,
β_k = parameter of x_{ntk},
z_{ns} = sth activity-generic attribute, and
κ_s = parameter of z_{ns}.

Representing Influence of Skewness Parameter

As explained, the parameter α describes the skewness of the response curve. It is natural to assume that the skewness may be different across individuals. In other words, some individuals may show highest sensitivity to change at $P_{nt}(y_{nt} = 1) = 0.5$, some at $P_{nt}(y_{nt} = 1) < 0.5$, and some at $P_{nt}(y_{nt} = 1) > 0.5$. However, it is difficult for an analyst to figure this out in advance. To accommodate such heterogeneity, this study therefore defines α as a function of some explanatory variables (z_{nq}) (e.g., individual attributes, travel cost, or travel party), where θ_q is the parameter of the qth variable z_{nq} and π is a constant term.

$$\alpha_n = \exp\left(\pi + \sum_q \theta_q z_{nq}\right) \tag{12}$$

The exponential function is adopted to meet the requirement that $\alpha_n > 0$. In the empirical analysis shown later, the scobit model with heterogeneous α (i.e., Equation 11) will be compared with the model having homogeneous α.

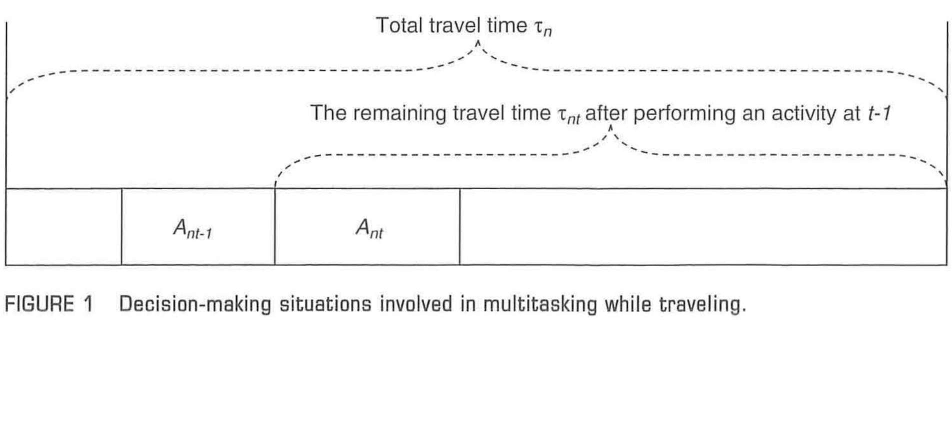


FIGURE 1 Decision-making situations involved in multitasking while traveling.

Estimation Method

With introduction of all the mentioned equations into the scobit model along the axis of travel time, the log likelihood function (LogL) can be obtained as follows:

$$\text{LogL} = \sum_{n=1}^{N} \sum_{t=1}^{T_n} \ln\left(P_t(y_t = 1)^{\delta_{nt}} P_t(y_t = 0)^{1-\delta_{nt}}\right) \tag{13}$$

where

N = total number of respondents,
T_n = number of time periods (i.e., number of activities performed during travel), and
δ_{nt} = dummy variable, which is equal to 1 when an activity is performed at wave t and otherwise 0.

The resulting model (Equation 12) is called the scobit-based panel model in this study, and it can be estimated by using standard maximum likelihood estimation methods.

DATA

A questionnaire survey was conducted in Hiroshima City, Japan, in December 2008. A total of 3,000 questionnaires were distributed at two major stations and a bus center on Tuesday, December 2; Thursday, December 4; and Sunday, December 7, 2008. To capture behavioral variations over the course of a day, the survey periods chosen were 07:00 to 09:00, 11:00 to 13:00, and 17:00 to 19:00. Respondents were asked to mail their completed questionnaires. Of the 679 questionnaires returned, 548 were valid. Of these, 215 reported multitasking behavior on the weekend (120 at 07:00 to 09:00, 56 at 11:00 to 13:00, and 39 at 17:00 to 19:00) and 333 on weekdays (187 at 07:00 to 09:00, 77 at 11:00 to 13:00, and 69 at 17:00 to 19:00). Respondents were asked to report their multitasking behavior during their use of public transportation systems at the time they received the questionnaires. In total, 2,246 activities were reported. For trip purpose, 51% of respondents were commuting, 12% were on business, 8% were going home, 8% were shopping, 6% were traveling for leisure, 3% were going to a hospital, 3% were going to school, and 9% were traveling for other reasons.

For the model analysis in this study, 523 valid samples were obtained by excluding some missing data. As shown in Figure 2, most of respondents carried out two or more activities during travel:

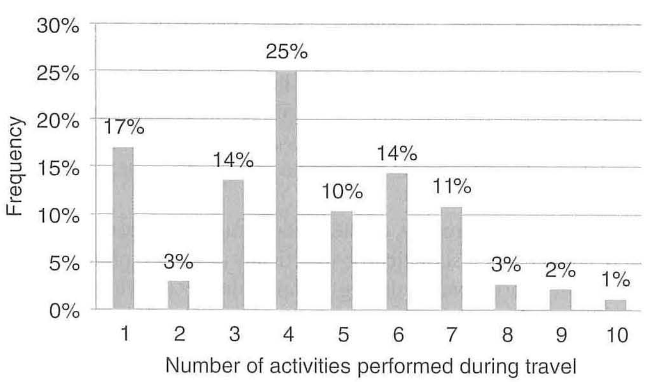


FIGURE 2 Distribution of number of activities performed during travel.

25% of respondents conducted four activities, and 10% to 14% participated in three, five, six, or seven activities, respectively.

Further, 31% of the activities performed during travel were window gazing or people watching, 18% were unknown activities, 17% involved doing nothing, 7% talking, 7% reading, 5% music listening, and 5% sleeping (Table 1). About 4% of activities are multidimensional, for example, reading a book while listening to music. Here, reading, talking, sleeping, listening to music, e-mailing via cell phone, web surfing via cell phone, doing business-related activities, and looking at advertisement posters are treated as activity participation, whereas window gazing or people watching, doing nothing, and other activities are treated as nonactivity participation. Although 18% of the activities were related to the other types of activities, because the contents of these activities were unclear, they were regarded as nonactivity participation. In this sense, the analysis might underestimate the influence of multitasking during travel. As a result, it was shown that in about one-third of all time periods in this analysis, activities were clearly performed during travel.

The other characteristics of respondents and activities are also shown in Table 1. About 60% of respondents experienced a travel time of less than 1 h; 25.6%, between 1.0 and 1.5 h; and 17.0%, longer than 1.5 h. Also, 87.4% of respondents traveled alone, and more than 50% of activities were performed in the morning; 24.5%, around noon; and 19.1%, in the afternoon or the early evening. It was found that about 30% of respondents had a seat during travel.

MODEL ESTIMATION AND DISCUSSION

This study attempted to clarify how to properly represent activity generation during travel and to examine influential factors, especially travel time, affecting activity generation. This section shows the effectiveness of the developed scobit-based panel model and clarifies the influential factors. Then the influence of travel time on the activity generation is examined, with a focus on travel time elasticity.

Explanatory Variables and Activities

A total of 523 valid samples were adopted for the model analysis. Major characteristics of this data set are shown in Table 1 and were discussed in the previous section. The selection of explanatory variables for the scobit-based panel model is based on Equations 11 and 12. The candidates are shown in Table 2, including activity-generic and activity-specific attributes, where the former further includes individual attributes and travel-related attributes.

As shown in Figure 2, the number of activities performed during travel ranges from one to 10. In theory, the individuals with the maximal number 10 can also be used for model estimation. Since in that case the corresponding sample size for those individuals is too small, for the sake of stable model estimation, the maximum number was arbitrarily limited to six. Individuals who performed more than six activities were grouped. Grouping the activities is straightforward. The problem is how to define some of the activity-specific attributes. This definition is done as follows:

1. Travel by rail. The dummy is set to 1, as long as any activity after the sixth activity is performed during travel by rail; otherwise it is set to 0.

TABLE 1 Characteristics of Samples

Attribute	Percentage
Activity-Generic Attributes	
Sex	
Male	49.3
Female	50.7
Age, years	
<19	2.9
20–29	9.0
30–39	15.3
40–49	23.5
50–59	28.1
>60	21.2
Travel time: one-way (min.)	
<30	17.8
30–60	39.6
60–90	25.6
>90	17.0
Travel cost: one-way (yen)	
<500	66.3
500–1,000	16.3
1,000–1,500	3.6
1,500–2,000	3.4
>2,000	10.3
Travel party	
Alone	87.4
With spouse	2.4
With child(ren)	1.9
With other family member(s)	1.7
With friend(s)	3.3
With one's superior	0.3
With workmate(s)	1.6
With other person(s)	1.4
Activity-Specific Attributes	
Whether to take a seat or stand during travel	
Stand (in a relaxed position)	17.8
Stand (in a crowded position)	3.1
Seated (on a longer chair)	11.0
Seated (in a box-type seat: face-to-face)	21.5
Moving to/from station or bus stop	46.7
Day of week	
Weekday	61.8
Weekend	38.2
Time of day	
Morning	56.4
Noon	24.5
Afternoon–early evening	19.1
Activity participation	
Yes	
Talking	6.8
Reading	6.7
Music-listening	5.3
Sleeping	4.9
Mixed activities	3.0
E-mail via cell phone	2.5
Doing business task	1.8
CM poster viewing	1.2
Web-surfing via cell phone	1.1
No	
Window-gazing or people-watching	31.3
Unknown activities	17.8
Doing nothing	17.0
Mixed nonactivities	0.8

TABLE 2 Explanatory Variables Used in Model Analysis

Item	Description
Activity-Generic Attributes: z_{ns} in Equation 11 and z_{nq} in Equation 12	
Individual attributes	Age (actual age) Sex (1: male, 0: female)
Travel-related attributes	Total travel time Total travel cost "Day of week" dummy (1: weekday, 0: weekend) "Time of day" dummy 1) Morning dummy (1: yes, 0: no) 2) Noon dummy (1: yes, 0: no)
Activity-Specific Attributes: y_{nt-1}, τ_{nt}, and x_{ntk} in Equation 11	
Activity participation in the past time period: y_{nt-1}	Equal to 1 in case of participation and 0 in case of nonparticipation. It is used to represent the influence of state dependence.
Remaining travel time: τ_{nt}	Remaining travel time when deciding to perform an activity during travel.
Travel by rail (1: yes, 0: no) Inside vehicle (1: yes, 0: no) Travel alone (1: yes, 0: no) Taking a seat (1: yes, 0: no)	These are some representative situational variables related to the decision on activity participation.

2. Inside vehicle. The dummy is set to 1, as long as any activity after the sixth activity is performed inside a vehicle; otherwise it is set to 0.

3. Taking a seat. The dummy is set to 1, as long as any activity after the sixth activity is performed by taking a seat; otherwise it is set to 0.

Results of Model Estimation

First, the model was estimated with a homogeneous skewness parameter. That is, a single skewness parameter was estimated. The results are shown in Table 3. The χ^2 statistic is calculated to test the null hypothesis that all parameters of the explanatory variables (excluding the skewness parameter) are zero. It is equal to 714.92, which is much larger than the critical value of 14.07 (the degree of freedom is 7, i.e., the number of parameters including the skewness parameter) at the 95% significance level. The adjusted Rho-squared is 0.2585. These results suggest that the obtained model is acceptable. The skewness parameter is estimated to be 0.8011, which is significantly different from both 0 and 1. Since it is known that if the skewness parameter is equal to 1, the scobit model returns to the logit model, the estimation results mean that the probability of activity participation during travel follows the more-realistic and flexible scobit assumption rather than the logit assumption. In other words, the logit model is not applicable in this case study, and use of the logit model could lead to seriously wrong estimates of people's responses to policies.

The estimation results of the other parameters show that state dependence has a statistically significant influence on the utility of activity participation during travel at the 95% level. Since the parameter of state dependence is positive, behavioral inertia appears to exist: that is, on average, participation in an activity in the past leads to participation in another activity during the remaining travel time. Such significant state dependence also implies that people

TABLE 3 Estimation Results: Model with Homogeneous Skewness Parameter

Parameter	Estimated Parameter	t-Score
Skewness of response distribution (α)	0.8011	9.619 (−2.389)
State dependence	0.9739	6.417
Remaining travel time	0.0018	1.675
Travel by rail (1: yes, 0: no)	0.2661	2.059
Inside vehicle (1: yes, 0: no)	1.0623	6.926
Travel party (1: travel alone, 0: otherwise)	−2.0752	−19.959
Taking a seat (1: yes, 0: no)	1.4331	9.159
Initial log likelihood[a]	−1,382.66	
Final log likelihood	−1,025.20	
Adjusted rho-squared	0.2585	
χ^2-statistic[b]	714.92 (>14.07[c])	
Sample size	523	

[a] α is fixed at the estimated value, and the other parameters are fixed at 0.
[b] χ^2-statistic = −2 * (initial log likelihood – final log likelihood).
[c] Critical value of χ^2-statistic with degree of freedom 7 (number of parameters) at the 95% significance level.

TABLE 4 Estimation Results: Model with Heterogeneous Skewness Parameter

Parameter	Estimated Parameter	t-Score
Skewness of response distribution: α (exponential function)		
Age (actual age)	−0.0083	−4.450
Sex (1: male, 0: female)	0.0521	0.703
Travel cost	−0.0028	−0.235
Date dummy (1: weekday, 0: weekend)	0.1254	1.829
Morning dummy (1: yes, 0: no)	−0.0810	−0.957
Noon dummy (1: yes, 0: no)	0.0108	0.115
State dependence	0.9937	7.245
Remaining travel time	0.0023	1.975
Travel by rail (1: yes, 0: no)	0.2649	1.974
Inside vehicle (1: yes, 0: no)	1.0792	6.742
Travel alone (1: yes, 0: no)	−2.0147	−17.797
Taking a seat (1: yes, 0: no)	1.5388	9.651
Initial log likelihood[a]	−1,459.77	
Final log likelihood	−1,016.47	
Rho-squared	0.3037	
χ^2-statistic[b]	886.60 (>21.03[c])	
Sample size	523	

[a] All the parameters are fixed at 0.
[b] χ^2-statistic = −2 * (initial log likelihood – final log likelihood).
[c] Critical value of χ^2-statistic with degree of freedom 12 (number of parameters) at the 95% significance level.

prefer to do some activities during travel. Since it is clear that in about one-third of the time periods in this case, study activities were performed during travel, the possibility that a trip has a positive utility would not be low. The other statistically significant factors are travel by rail or not, inside vehicle, travel alone or not, and taking a seat or not. Those trip makers who traveled by rail, were inside a vehicle (not at a platform or bus stop), traveled with other persons, or had a seat during travel tended to perform some activities. Intuitively, this is logical. The remaining travel time is estimated to be positive, meaning that the longer the travel time, the higher the probability to perform some activities. Although its parameter is statistically significant just at the 90% level, this result suggests that trip makers could derive some positive utility from use of travel time. This does not mean that performing activities during the whole travel time will automatically result in positive utility of travel. It also does not necessarily mean that at some very long distance, the utility of travel could become positive. Rather, the sign of the total utility derived during travel is determined by the trade-off between activity participation during travel and trip making itself.

Next, to further improve model performance and confirm the influence of travel time on activity participation during travel, the scobit-based panel model with a heterogeneous skewness parameter is estimated. The skewness parameter is defined as an exponential function to meet the condition that $\alpha > 0$. Although, in theory, activity-generic attributes could be used to explore both activity participation and the skewness parameter, but here in the first attempt to apply the scobit model in transportation, to avoid the confusing interpretation of the influence of activity-generic attributes, they are introduced only into Equation 12, that is, the skewness parameter function. The estimation results are shown in Table 4. It is confirmed that model accuracy is significantly improved from 0.2585 to 0.3037. To further confirm the better performance of the model with a heterogeneous skewness parameter, the χ^2 statistic is calculated by using the respective final log likelihoods from the two models. The χ^2 statistic is equal to 17.46 [= −2 * (−1,025.20) − (−1,016.47)], which is larger than the critical value 11.07 (the degree of freedom is 5, i.e., the number of parameters newly introduced in the model with heterogeneous skewness parameter) at the 95% significance level. This suggests that the model with a heterogeneous skewness parameter is superior to the model with a homogeneous skewness parameter. The parameter for the remaining travel time was estimated to be significant at the 95% level, still with a positive sign. Regarding the model with a homogeneous skewness parameter, all parameters related to utility function are statistically significant at the 95% level. Concerning the parameters included in the skewness parameter function, only age significantly affects the skewness parameter and its parameter has a negative sign. This result restates the importance of introducing heterogeneity into the travel behavior model.

Elasticity of Travel Time

To examine how travel time influences activity participation during travel, travel time elasticity of the probability of activity participation was calculated on the basis of the model with a heterogeneous skewness parameter. The elasticity is as follows:

$$E_{x_{ntk}}^{P_t(y_t=1)} = \alpha_n \exp(v_t)\beta_k x_{ntk} \frac{P_{nt}(y_{nt}=0)}{P_{nt}(y_{nt}=1)(1+\exp(v_t))} \tag{14}$$

Calculation results for each activity (maximum = 6) are shown in Figure 3. It can be seen that the elasticity varies largely across individuals, and such variations gradually decrease with an increase in the number of activities. Specifically, the standard deviation largely decreases from $\sigma = 0.075$ in the first activity [Activity 1 (Figure 3*a*)] to $\sigma = 0.052$ in the third activity [Activity 3 (Figure 3*c*)] and becomes

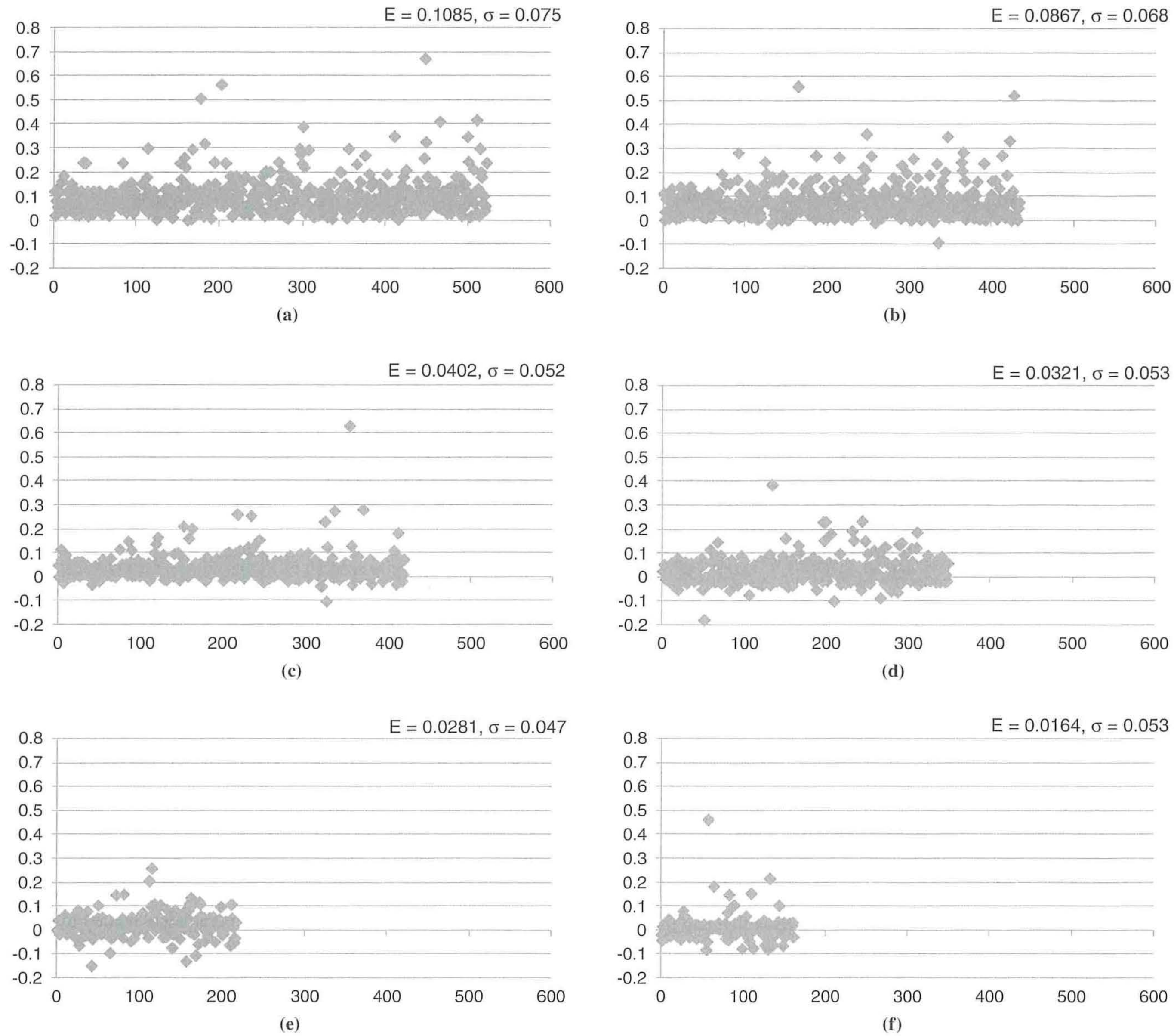


FIGURE 3 Distribution of travel time elasticity (horizontal axis: individuals; vertical axis: elasticity; E = average elasticity; σ = standard deviation).

almost invariant from the fourth activity (Figure 3*d*) at value 0.5. The average elasticity (calculated by using the average values of variables introduced into the model) first largely decreases from 0.1085 at the first activity (Figure 3*a*) to 0.0402 at the third activity (Figure 3*c*) and then gradually decreases to 0.0164 at the sixth activity (Figure 3*f*). These results suggest that the elasticity shows a dramatic change from the third activity onward. After the third activity, trip makers become more insensitive to change in travel time.

CONCLUSIONS AND DISCUSSION

Focusing on the multitasking behavior of public transportation users during travel, this study examined how to properly model activity participation to estimate the effects of influential factors. Data on multitasking behavior collected in Hiroshima City, Japan, in December 2008, contributed to the following:

1. In this case study, on average, public transportation users conducted four or more activities during a single trip, and in about one-third of the time periods under study, activities were clearly performed during travel.
2. A new model, the scobit-based panel model, was developed to represent repeated participation in activities during the course of travel. Unlike the popular logit or probit models, the scobit model does not assume that individuals with a probability of 0.5 for choosing to perform an activity are most sensitive to changes in explanatory variables. Model estimation results confirmed that the skewness parameter is significantly different from 1, implying that the scobit model is more appropriate for describing activity participation choice

during travel than is the logit model. Furthermore, introducing a heterogeneous skewness parameter is more effective for representing activity participation during travel than is assuming a homogeneous skewness parameter across individuals.

3. The estimated state dependence parameter introduced into the scobit-based panel model suggests that people prefer to repeatedly perform some activities during travel. This implies that travel may generate some positive utility.

4. Model estimation results reveal that length of travel time is proportional to the probability of activity participation during travel. Calculation results of travel time elasticity show that activity participation is sensitive to change in the travel time up to the first half of travel time and becomes less sensitive to travel time later.

This study was motivated by the desire to explore whether the utility of travel time is necessarily negative. Although it was not directly examined whether the utility of travel time is negative, it has been clarified that activity participation is preferred during travel, and its probability is proportional to the length of travel time. However, there remain some unresolved issues. For example, it is necessary to make clear how much activity participation during travel could mitigate negative feelings of travel and consequently reduce the negative utility of travel. This should be done by properly representing the trade-off between activity participation and trip making itself. Such trade-off may differ across different types of activities, across different travel modes, and between short-distance and long-distance trips. These research questions should be addressed in future research. Especially, it is important how to give a more convincing and more accountable valuation of travel time savings by reflecting the influence of multitasking during travel. Related to this discussion, it could be important to look at what a positive utility of travel time means from a behavioral perspective. The day reconstruction method proposed by Kahneman et al. (*15*) may be an interesting candidate approach in this context because it can measure whether performing an activity leads to positive or negative feelings. Activity participation and duration should be represented simultaneously by, for example, linking the scobit modeling technique with a time allocation model. In addition, it is worth examining how multitasking while traveling might be interrelated with other activities conducted at stationary locations (e.g., home or office). To enhance model accountability, it should be clarified how to specify explanatory variables (both activity-generic and activity-specific variables) to express activity participation and duration. More studies about the multitasking linking it to ICTs, value of activity time, and policies to further promote the use of public transportation systems should be conducted.

REFERENCES

1. Kenyon, S., and G. Lyons. Introducing Multitasking to the Study of Travel and ICT: Examining Its Extent and Assessing Its Potential Importance. *Transportation Research Part A,* Vol. 41, No. 2, 2007, pp. 161–175.
2. Ettema, D., and L. Verschuren. Multitasking and Value of Travel Time Savings. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2010,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 19–25.
3. Timmermans, H. J. P., and P. J. H. J. van der Waerden. Synchronicity of Activity Engagement and Travel in Time and Space: Descriptors and Correlates of Field Observations. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 1–9.
4. Faulks, I., J. D. Irwin, and E. Chekaluk. Differences in Response of Male and Female Drivers to Everyday Distractions. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
5. Timmermans, H., and J. Zhang. Analyses of Multi-Tasking on Public Buses During Work Commute Using Support Vector Machines and Boosted Regression Trees. *Proc., 14th International Conference of Hong Kong Society for Transportation Studies,* 2009.
6. Mokhtarian, P. L., I. Salomon, and S. L. Handy. The Impacts of ICT on Leisure Activities and Travel: A Conceptual Exploration. *Transportation,* Vol. 33, No. 3, 2006, pp. 263–289.
7. Schwanen, T., and M. P. Kwan. The Internet, Mobile Phone and Space-Time Constraints. *Geoforum,* Vol. 39, No. 3, 2008, pp. 1362–1377.
8. Sheller, M., and J. Urry. The New Mobilities Paradigm. *Environment and Planning A,* Vol. 38, No. 2, 2006, pp. 207–226.
9. Kenyon, S. The "Accessibility Diary": Discussing a New Methodological Approach to Understand the Impact of Internet Use upon Personal Travel and Activity Participation. *Journal of Transport Geography,* Vol. 14, No. 2, 2006, pp. 123–134.
10. Jain, J., and G. Lyons. The Gift of Travel Time. *Journal of Transport Geography,* Vol. 16, No. 2, 2008, pp. 81–89.
11. van der Waerden, P. J. H. J., H. J. P. Timmermans, and R. van Neerven. Extent, Nature, and Covariates of Multitasking of Rail Passengers in an Urban Corridor: A Dutch Case Study. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2110,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 106–111.
12. Nagler, J. Scobit: An Alternative Estimator to Logit and Probit. *American Journal of Political Science,* Vol. 38, 1994, pp. 230–255.
13. Golob, T. F., R. Kitamura, and L. Long. *Panels for Transportation Planning,* Kluwer Academic Publishers, Dordrecht, Netherlands, 1997.
14. Burr, I. W. Cumulative Frequency Functions. *Annals of Mathematical Statistics,* Vol. 13, 1942, pp. 215–232.
15. Kahneman, D., A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science,* Vol. 306, No. 5702, 2004, pp. 1776–1780.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Eliciting the Needs That Underlie Activity–Travel Patterns and Their Covariance Structure

Results of Multimethod Analyses

Linda Nijland, Theo Arentze, and Harry Timmermans

The modeling of dynamic activity generation is high on the research agenda in activity-based transport demand modeling. The concept of dynamic needs has been put forward as such a mechanism. Needs that underlie the generation of such discretionary activities as social, recreational, and sports activities is investigated. Three surveys were conducted to elicit, establish, and analyze the needs. Qualitative face-to-face interviews were carried out with a laddering technique to reveal need dimensions through an exhaustive classification of discretionary activities. Quantitative approaches were then used to determine which needs are equivalent in their effects on activities and, hence, can be merged. Finally, a questionnaire-based survey involving a large sample of individuals was used to measure personal levels of the needs identified and to correlate these measures with socioeconomic and behavioral characteristics. Six independent needs emerged from this research: physical exercise, social contact, relaxation, fresh air and outdoors, new experiences, and entertainment. Many-to-many relationships between activities and needs support the hypothesis that substitution relationships may play a significant role in activity generation. This observation implies that current practice in activity-based modeling of focusing on activities may produce biased results in development of dynamic models of transport demand. Furthermore, the results show that personal levels of these needs correlate with various socioeconomic as well as behavioral variables.

Although progress in activity-based models has been formidable and these models are now slowly but gradually moving to practice, there is still ample room for improvement. An issue requiring elaboration and further attention concerns the classification of activities. Several existing activity-based models are based on a simple classification of mandatory and discretionary activities (sometimes differentiating maintenance and social or leisure). Empirical results, however, indicate that these models perform better for work and shopping activities than for social, recreational, and leisure activities. In part, this may be because the motivators underlying these latter activities are more varied and because the choice options are both larger in number and more diverse and hence are more difficult to predict. However, relatively poor results may reflect that these activities are partly substitutable because they satisfy common underlying needs. For example, both shopping and socializing mean a break from housekeeping duties. Shopping will also contain an element of meeting other people and hence will partly satisfy some general social needs.

Doherty (*1*) and Doherty and Mohammadian (*2*) also discussed the issue of classification of activities, albeit from a different perspective. Examining planning horizons, they found evidence that the process of planning activities is not congruent with commonly assumed hierarchical processes in activity-based models. They applied an ordered probit model to analyze the influence of a series of factors. Closely related to the problem of classification is the issue of activity generation. Mechanisms underlying activity generation are still poorly understood and not well represented in current activity-based models (*3, 4*). The notion that daily activities of individuals are driven by basic needs has been at the core of the activity-based approach since the pioneering work of Chapin (*5*) and is further emphasized by Miller (*6*) and Axhausen (*7*). Miller derived some elements of his framework for modeling short- and long-term household-based decision making from Maslow's hierarchy of needs. Meister et al. partially implemented needs into their operational model of activity scheduling (*8*).

To incorporate possible substitutions between activities, Arentze and Timmermans developed a need-based model (*9*). They defined the utility of an activity by its contribution to the satisfaction of dynamically changing needs. So-called potentials describe relationships between activities and needs quantitatively. Potentials depend on the nature of the activity and on attributes such as duration, location, and time of day. The model predicts the timing and duration of activities in a dynamic longitudinal framework, taking into account time budget constraints and needs at both household and person levels. The results of numerical simulations supported the face validity of the suggested theoretical framework and modeling approach and demonstrate the possibility of incorporating substitution effects between activities and complex dynamic interactions between activities in general. To date, however, their approach lacks empirical validation.

The purpose of this study is to fill that gap. A good classification of needs that underlie activity generation is sought, and the nature of the relationships between activities and underlying needs is examined. The design and results of three related surveys carried out to elicit, establish, and analyze the needs underlying activity–travel patterns of individuals are described. The first survey uses qualitative face-

Faculty of Architecture, Building, and Planning, Urban Planning Group, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands. Corresponding author: L. Nijland, e.w.l.nijland@bwk.tue.nl.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 54–62.
DOI: 10.3141/2157-07

to-face interviews to find out which needs and other factors are responsible for the discretionary activities individuals conduct in daily life. This resulted in nine needs to be included in additional research. In the second survey, subjects were asked to indicate to what extent they think the needs are influenced by 22 types of recreational, social, and sports activities. After looking at the similarities between the needs and their influences on performing activities, the set could be reduced to six independent needs. The final survey is part of a larger questionnaire. In this survey, statements were use to measure, by means of scale construction, the general levels of these needs of the respondents and analyze the correlations of the levels with relevant socioeconomic and behavioral characteristics.

SURVEY 1. IDENTIFYING NEEDS

The first survey was done to identify which needs play a role in the generation of discretionary activities. To elicit the needs and other factors underlying activity choice, qualitative face-to-face interviews were carried out. In the interviews, an exhaustive set of 22 social, leisure, and sports activities were taken into account (see Figure 1). This set was identified from existing activity-diary data. Although a qualitative approach was used, the interview had a fixed structure. It began with a question asking which activities the respondent never or rarely conducts. Only the activities that the subject conducts on a regular basis were included in the interview. Activity choice sets

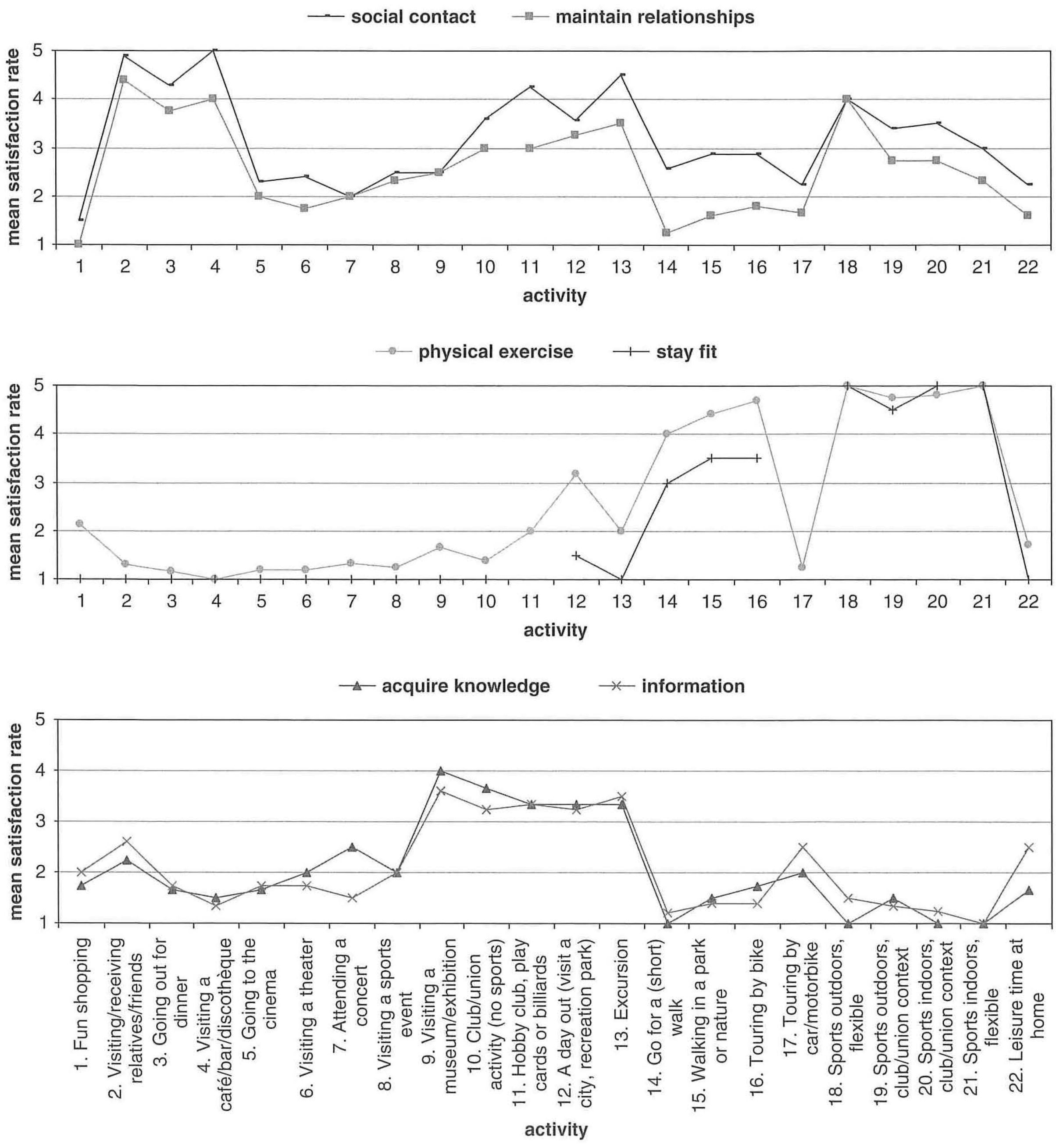


FIGURE 1 Comparison of mean satisfaction rates, Survey 1.

were generated by randomly selecting combinations of three activities such that all activities appear at least once in a choice set. In practice, this meant that the number of hypothetical choice sets varied between three and six, dependent on the number of activities the respondent conducts on a regular basis. For each choice set, a hypothetical scenario was presented of the following general form: "Assume that on a day there is time available to conduct an activity and that you can choose between the following three activities." Then the three randomly selected activities were shown.

A laddering technique (*10*) following logic similar to that of the so-called CNET method (*11, 12*) was used to elicit the factors underlying activity choice. The following question was asked: "What are your considerations when choosing between the three activities?" If a consideration was not clear or could not be identified as a need, the "Why is that important to you?" question, which is typical for laddering, was posed.

This method was used to get a broader view of the factors, not restricted to needs, that influence activity choice. Because this method is exploratory and free of theory, it is possible that unexpected needs or other factors occur. For every factor that does not correspond to a need, respondents were asked to indicate the importance of that factor for making the decision (0 = not important, 1 = somewhat important, 2 = important, and 3 = very important). Furthermore, to obtain a first indication of the relationships between activities and needs, the subjects were asked to indicate, for each of the 22 activities they conduct on a regular basis, to what extent, on a five-point scale, the activities satisfy the needs they mentioned (1 = not at all, 5 = to a large extent).

Survey 1 Sample

Eight people were interviewed. After the eight interviews, it was decided that no more were needed for this phase. Convergence of responses suggested that the outcome would not be different if additional persons were interviewed. Although this is typically a convenience sample, care was taken to have a representation of diversity. Individuals from different stages of life and environments were included. Four respondents lived in a city and the other four in a village. Three respondents had a full-time job, one was a full-time student, another worked part time, and the remaining three had a lot of discretionary time because they were either retired or looking for work. Two of the subjects had young children and two others have children who had left home. Ages ranged from 28 to 64 years, and 50% were female.

Survey 1 Results

A distinction was made between considerations that could be described as needs and other factors (e.g., weather, time of day).

Factors

Table 1 shows all factors that were indicated by the respondents, how often they were mentioned, and the mean importance for making the decision (0 = not important, 1 = somewhat important, 2 = important, and 3 = very important). The main factors that came out of the interviews were weather, time of day, duration of the activity and available time, day of the week, and if there is a particular reason for conducting the activity (e.g., need to buy something or there is something happening). The distribution of mean importance scores shows that the factors that are relatively important include goods or offerings, social obligation, there is something happening, need to buy something, and dependent on the availability of others.

TABLE 1 Factors Influencing Choice of Activity

Factor	Respondents	Total Frequency	Mean Importance
Weather	7	19	1.78
Time of day	6	16	1.73
Activity duration or available time	6	8	1.25
Need to buy something	6	7	2.1
There is something happening	5	12	2.2
Day of the week	4	4	1.67
Dependent on the availability of others	3	10	2.08
Costs	3	9	1.67
Time elapsed since the previous performance	3	9	1.3
Range of goods (offerings)	3	7	2.33
Time of year	3	3	1.5
Combining activities	3	3	0.84
Social obligation	2	4	2.25
Time pressure	2	3	1.25
Available transport modes	2	2	2
With whom	2	2	2
Distance	2	2	1.5
Opening hours	2	2	1
Mood	1	2	2
Level of fitness	1	2	1.5
Habit	1	1	1
Entertainment value	1	1	1

Needs

This paper focused on needs. Table 2 shows the needs mentioned by subjects and how often the needs were mentioned in the hypothetical scenarios. The needs for social contact, relaxation, and physical exercise were indicated by nearly all respondents. Other frequently occurring needs were fresh air and being outdoors, maintaining relationships, and new experiences. Furthermore, 50% of the subjects mentioned the need for information, nature, acquiring knowledge, rest, and entertainment. Respondents with young children added the need to guide their children's development.

Impact of Activities on Needs

The results were used to determine if some needs are basically the same and can be combined or if one of them can be deleted. The average satisfaction rates across the activities can be determined to indicate an activity profile of a given need. For activity generation, any two needs that have the same activity profile can be merged into a single need category, because they are equivalent and cannot be

TABLE 2 Needs Influencing Choice of Activity

Need	Respondents	Total Frequency
Social contact	8	25
Relaxation	8	24
Physical exercise	7	20
Fresh air and being outdoors	5	9
Maintain relationships	5	8
New experiences	5	8
Information	4	8
Nature	4	7
Acquire knowledge	4	6
Rest	4	5
Entertainment	4	5
Variety	3	5
Personal development	2	5
Curiosity	2	4
Going out for a short time	2	4
Stay fit	2	4
Guide your child's development	2	2
Flexibility	1	1
Be occupied in a creative way	1	1
Adventure	1	1
Freedom	1	1
Buying something new	1	1
Make a social contribution	1	1
Reflection	1	1
Clear your mind	1	1
Luxury	1	1
Competition	1	1
Enjoyment	1	1
Being there for other people	1	1

distinguished. Figure 1 shows the activity profiles for the pairs of needs social contact and maintaining relationships, physical exercise and staying fit, and acquiring knowledge and information. On the basis of these results, the following pairs of needs can be merged, because they show similar activity profiles:

- Maintaining relationships and social contact,
- Staying fit and physical exercise, and
- Acquiring information and knowledge.

For the other needs, the similarities were not as clear. Thus, nine important needs remained. They were mentioned by at least half the respondents and include the following:

- Social contact, relaxation, and physical exercise (100% respondents mentioning);
- Fresh air and being outdoors, new experiences, and information or acquiring knowledge (63%); and
- Nature, rest, and entertainment (50%).

Combinations of these dimensions showed considerable overlap for their attribute profile. Whether these overlaps warrant a further reduction of the set could not be established with certainty because of the small sample size. Therefore, a second round using a larger sample was implemented to determine whether the set could be further reduced.

SURVEY 2. ESTABLISHING NEEDS

The next phase of data collection verified the independence of the needs identified in the previous phase. Similar to the approach used in the previous section, subjects were asked to indicate, for each of the nine remaining needs and for the same list of 22 social, leisure, and sports activities (Figure 1), to what extent the activity satisfies the need. An Internet questionnaire was developed to collect the data for a larger sample. The first part of the questionnaire consisted of general questions about socioeconomic characteristics. The second part showed for one need per page the list of activities and contained for every activity a field where the satisfaction rate could be entered. Respondents could indicate n/a (not applicable) if they did not conduct the activity regularly.

Survey 2 Sample

For this Internet questionnaire, about 45 acquaintances, students, and colleagues were approached by e-mail. The total sample consists of 41 persons. Three did not complete the questionnaire. The responses of the eight persons of the preceding face-to-face interviews were included; however, they filled out only the satisfaction rates for the needs they mentioned themselves. Table 3 describes the sample for distribution on some relevant socioeconomic variables, and for comparison the table shows the same distributions in the population at the national level. The table shows that the elderly (65+) and single-adult households are underrepresented and that above-average educated groups are overrepresented.

Survey 2 Results

For every activity and each need, the average satisfaction rate was calculated, resulting in an activity profile for each need. The results

TABLE 3 Composition of Samples

	Sample Survey 2 (%)	Sample Survey 3 (%)	Population (%)
Gender			
Female	51	53	50.5
Male	49	47	49.5
Age (years)			
$15 \leq 25$	12	7	15
$25 \leq 45$	58	48	37
$45 \leq 65$	25	34	33
$65 \leq 85$	5	10	16
Education			
Below average	8	14	35
Average	20	25	41
Above average	72	61	24
Household composition			
Single, no children	15	23	35
Single, children	0	3	6
Double, no children	54	38	29
Double, children	29	33	29
Multiple persons	2	1	1

are shown in Figure 2. The activity profiles of social contact and physical exercise clearly deviate from those of the other needs. The need to rest follows the line of the need for relaxation. Apart from a difference in scale, rest has the same influence and therefore can be subsumed under relaxation. Moreover, relaxation was mentioned by all subjects in the face-to-face interviews and rest by only half the subjects. Thus, a need to rest cannot be distinguished from a need to relax when in activity generation and, therefore, these two needs can be merged into a single category. Figure 2 also shows that the activity profiles of fresh air and being outdoors and nature also are comparable. In the first survey, fresh air and being outdoors was mentioned more often than nature, so it was decided to remove nature from consideration. The activity profiles for acquiring knowledge and new experiences are strongly alike as well. The satisfaction rates of acquiring knowledge exceed only four times the level of three units. Therefore, acquiring knowledge

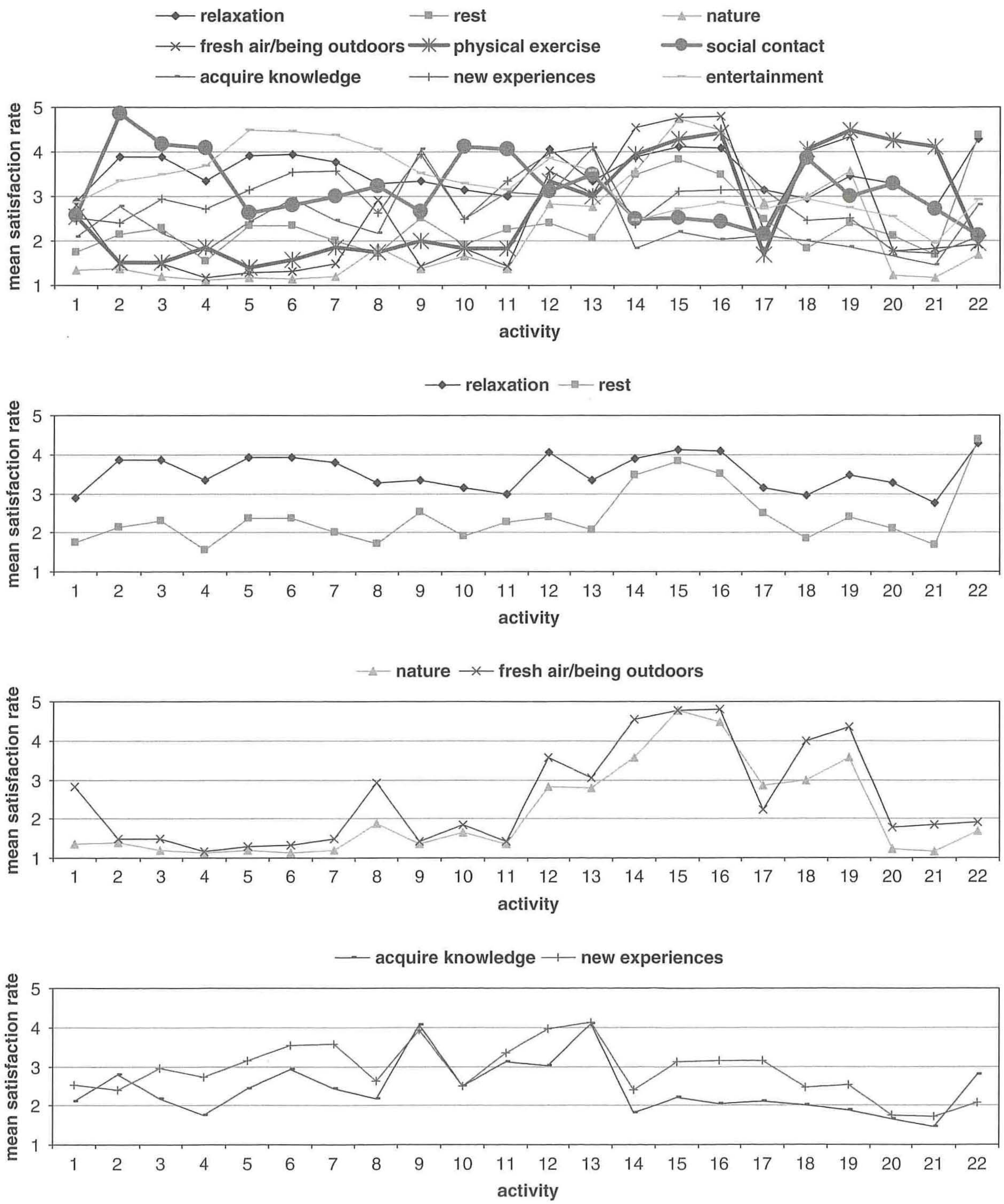


FIGURE 2 Comparison of mean satisfaction rates, Survey 2 (activities as shown in Figure 1).

is subsumed under new experiences. In the end, the following six needs remained and were included in the third-round survey:

- Social contact,
- Physical exercise,
- Relaxation,
- Fresh air and being outdoors,
- New experiences, and
- Entertainment.

SURVEY 3. ANALYZING NEEDS

The third round was done to measure individuals' basic levels of the needs identified in the previous rounds and to determine the extent to which these levels correlate with socioeconomic and behavioral characteristics. In this main survey, a much larger sample of individuals is used, and the six remaining needs are incorporated: physical exercise, social contact, relaxation, fresh air and being outdoors, new experiences, and entertainment. To construct a scale for each need, four statements were included in the questionnaire as indicators of the need: two were positively oriented and two negatively oriented. The statements generally start with "I think it is important to," "I like to," and "I have hardly any need for." A complete list of the statements is given in Table 4. The statements were mixed so that those related to one need were not shown next to each other. By using Likert scales, subjects indicated to what extent they agreed with the statements (totally disagree, disagree, neutral, agree, or totally agree). The scales were part of a larger questionnaire consisting of six parts. To validate the scales, the study focused on three parts: the statements, socioeconomic and demographic variables, and the standard week pattern of the respondents. In the standard week pattern, respondents indicated, for every day of the week, which of the given activities they normally (phrased as "almost always") conduct on that day. For each selected activity, the subjects were to specify the usual duration and travel time. This part included 18 activities, including work, education, bring or collect child(ren), grocery shopping, and some sports, leisure, and social activities. The decision to use the latter was based on frequencies of those activities indicated by respondents in the Amadeus survey (*13*). If only a very small percentage of the 1,600 respondents conducted the activity at least once a week, the activity was not included.

TABLE 4 Survey 3 Statements

Category	Statement
Physical exercise	1. I think it is important to practice a lot of sports. 2. I am not fond of doing sports. 3. I like to exercise a lot. 4. I have hardly any need for physical exercise.
Fresh air and outdoors	1. I prefer to travel by bike or by foot. 2. I prefer to stay indoors as much as possible. 3. I think it is important to do activities outdoors. 4. I have hardly any need for fresh air.
Social contact	1. I like to have people around me. 2. I have hardly any need for social contacts. 3. I think social contacts are important. 4. I like to be alone.
Relaxation	1. I think time for myself is very important. 2. I have hardly any need for relaxation. 3. I think leisure time is very important. 4. I like to be busy.
Entertainment	1. I like liveliness around me. 2. I have hardly any need for entertainment. 3. I like to be entertained. 4. I would rather go for a walk or cycling than go to the movies, the theater, or a concert.
New experiences	1. I think it is important to gain new experiences. 2. I avoid unfamiliar situations. 3. I am inquisitive in nature. 4. I have hardly any need to do new things.

Survey 3 Sample

Respondents were selected from a sample of neighborhoods in the Eindhoven region. In the last two weeks of June 2009, 4,000 invitation cards were distributed to households in the selected neighborhoods. Individuals who in an earlier survey had indicated their willingness to participate again in an Internet survey were contacted by e-mail (*14*). In this way, approximately 400 additional individuals were invited to participate in the survey. As an incentive, 20 vouchers worth 50 euros were allocated to respondents through a lottery. In total, 408 individuals started and 270 completed the main questionnaire.

Table 3 describes the sample and the Dutch national population with regard to some relevant socioeconomic variables. The sample is reasonably representative, except that above-average educated groups are overrepresented. This bias is typical for Internet surveys in general (*15, 16*). The elderly (65+ years) and young persons (<25 years) are somewhat underrepresented, and households consisting of two persons (married or living together) are somewhat overrepresented.

Survey 3 Results

The 270 respondents who completed the main questionnaire were included in the scale analysis. The four statements for every need, as shown in Table 4, were used to construct a scale for each need. To determine the reliability of each scale, Cronbach's alpha was calculated. A factor analysis was carried out to check whether the scale was homogeneous, that is, relates to a single dimension. The factor analysis was conducted only to check whether a single factor can explain shared variance for the four statements that were included to represent the need, rather than to identify the joint factors (needs).

The results (Table 5) show that the indicators for the needs physical exercise, social contact, new experiences, fresh air and being outdoors, and entertainment have a single dimensionality (i.e., only one component has an eigenvalue >1). The statements of the need for relaxation are not homogeneous according to this criterion. The first section of the table shows the Cronbach's alpha values. The "alpha without" value for each item shows the alpha value that would result if the item were removed from the set. Bold print indicates cases in which alpha without is larger than the current value. The Cronbach's alpha values suggest that the indicators for physical exercise, social contact, and new experiences are reliable to a satisfactory extent. For some needs, the alpha can be increased by deleting one statement (the value alpha without is higher than the overall alpha for the need). However, three statements is considered an absolute minimum for

TABLE 5 Analysis Results

	Physical Exercise	Social Contact	Relaxation	Fresh Air and Being Outdoors	New Experiences	Entertainment
Cronbach's Alpha						
Alpha	0.834	0.748	0.287	0.526	0.692	0.602
Alpha without						
Statement 1	0.758	0.652	0.018	**0.570**	0.564	0.587
Statement 2	0.783	0.623	0.102	0.311	0.671	0.444
Statement 3	0.808	0.643	**0.297**	0.406	0.667	0.457
Statement 4	0.806	**0.812**	**0.460**	0.514	0.602	**0.623**
Factor Analysis						
Eigenvalues						
Component 1	**2.700**	**2.307**	**1.525**	**1.748**	**2.104**	**1.882**
Component 2	0.542	0.853	**1.026**	0.984	0.795	0.916
Component 3	0.495	0.486	0.902	0.720	0.606	0.648
Component 4	0.263	0.354	0.547	0.548	0.495	0.554
New eigenvalues						
Component 1		**2.152**	**1.525**	**1.626**		**1.708**
Component 2		0.488	0.921	0.817		0.735
Component 3		0.360	0.555	0.557		0.557

NOTE: Bold numbers represent cases where the alpha without value is larger than the current value.

constructing a scale. In the cases of social contact, fresh air and being outdoors, and entertainment, one indicator (based on the level of Cronbach's alpha if item deleted) was deleted to increase reliability. For relaxation as well, one of the statements was left out. Nevertheless, the Cronbach's alpha for this need is still too low (an alpha value of 0.70 or larger is generally considered satisfactory). For social contact, Statement 4 ("I like to be alone") was deleted. For fresh air and being outdoors, Statement 1 ("I prefer to travel by bike or by foot") was excluded from further analyses. Entertainment Statement 4 ("I rather go for a walk or cycling than going to the movies, the theatre, or a concert") was left out of the final score variable, and for relaxation, Statement 4 ("I like to be busy") was deleted.

A new variable for each need can be constructed by summation of the scores of the indicators, where totally disagree is counted as 1, disagree as 2, neutral as 3, agree as 4, and totally agree as 5. Consequently, the summary variables have a minimum of 3 (if one indicator was deleted) or 4 (if all four statements were included) and a maximum of 15 or 20.

Correlation coefficients were calculated to determine whether there are correlations between the sum scores of the needs, on the one hand; and some socioeconomic variables and time spent on activities in the standard week pattern, on the other hand. Nominal and some ordinal variables were dummy coded. The variables of the standard week pattern were computed by taking the sum of the durations of each time the activity was selected. If the activity was not selected in the week pattern, the value 0 was assigned. Thus, the variables represent the total amount of time normally spent by the respondent on the activities every week and whether they conduct the activity on a weekly basis. Table 6 shows how the needs are related to the other needs and subsequently to the time spent on several activities in the week pattern and to other general variables (e.g., gender or age). The correlation coefficients (Spearman's rho) that are significant at the 0.01 level are identified by two asterisks, and those significant at the 0.05 level are identified with a single asterisk. Values higher than 0.2 are shown in bold. Some variables that did not show significant or high correlation coefficients were not included in the table, for example, household composition, car ownership or availability, age of youngest child, and dwelling types (other than detached).

The results show that several needs are somewhat correlated. Especially, the need for social contact is related to entertainment, new experiences, and fresh air and being outdoors. Furthermore, the need for fresh air and being outdoors is correlated to the need for physical exercise. In all cases, the needs are positively interrelated; this interrelation means that in general some subjects tend to give high scores for all needs and, conversely, others indicate relatively low rates.

The standard week pattern indicates that if the amount of hours a week spent on paid work and education is higher, the need for physical exercise, new experiences, and entertainment increases. The need for social contact is higher when respondents spend more time on visiting or receiving relatives and friends and visiting a café, bar, or discothèque. For students, the need for relaxation grows when time spent on doing homework for school and study increases. Persons who exercise frequently show a higher need for physical exercise, entertainment, and fresh air and being outdoors. Finally, the need for entertainment is somewhat positively correlated to the time spent on visiting a café, bar, or discothèque, visiting or receiving relatives and friends, and fun shopping.

In addition to the behavioral variables, socioeconomic factors have several effects as well. Older age groups display lower needs in general. In particular, the needs for social contact and entertainment decrease as people age. Similarly, retired subjects have a lower need for entertainment. If the number of children in the household is greater, the need for fresh air and being outdoors is lower, and the more often individuals work from home, the less time they need for relaxation. Furthermore, females tend to have higher needs for physical exercise, new experiences, and entertainment. Respondents living in a detached house display a higher need for physical exercise and fresh air and being outdoors. Finally, the need for fresh air is higher for those who have a driver's license.

TABLE 6 Correlation Coefficients (Spearman's Rho)

Variable	Physical Exercise	Fresh Air and Outdoors	New Experiences	Social Contact	Relaxation	Entertainment
Needs						
Physical exercise	1.000	**0.468****	**0.244****	**0.318****	0.166**	**0.304****
Fresh air and being outdoors	**0.468****	1.000	**0.374****	**0.426****	0.185**	0.166**
New experiences	**0.244****	**0.374****	1.000	**0.436****	**0.218****	**0.289****
Social contact	**0.318****	**0.426****	**0.436****	1.000	**0.330****	**0.463****
Relaxation	0.166**	0.185**	**0.218****	**0.330****	1.000	**0.227****
Entertainment	**0.304****	0.166**	**0.289****	**0.463****	**0.227****	1.000
Week pattern (hours per week)						
Paid work	0.178**	0.029	0.153*	0.119	0.086	0.169**
Work and education	**0.238****	0.044	0.194**	0.117	0.106	**0.235****
Homework	**0.221**	**0.381**	−0.033	−0.110	**0.516***	0.057
Physical exercise	**0.651****	**0.230****	0.161**	0.183**	0.168**	**0.273****
Walking or cycling	0.107	**0.273****	0.060	−0.034	0.023	−0.030
Café, bar, or disco	**0.211****	0.080	0.171**	**0.226****	0.082	**0.330****
Visiting or receiving relatives and friends	0.110	0.071	0.092	**0.245****	0.114	**0.303****
Fun shopping	−0.094	−0.008	0.119	0.182**	0.137*	**0.255****
General variables						
Gender (1 = female)	−0.171**	−0.015	0.087	0.118	0.093	0.060
Age	−0.118	0.066	−0.147*	**−0.208****	−0.179**	**−0.318****
Income	0.034	0.051	0.136*	−0.098	−0.023	−0.022
Education level	0.072	−0.044	0.073	−0.136*	−0.013	−0.065
Dominant activity						
Paid work	0.014	0.014	0.120	0.035	0.125*	0.103
Education or study	0.012	−0.085	0.018	0.069	0.057	0.126*
Retired	−0.028	0.040	−0.098	−0.070	−0.182**	**−0.227****
No. of children	0.130	**−0.225***	0.113	−0.055	0.026	0.059
Detached house	0.166**	0.124*	−0.085	−0.006	−0.014	−0.065
Driver's license	−0.005	0.183**	0.055	0.065	−0.013	0.015
Work from home	0.094	−0.046	−0.072	−0.013	**−0.255***	−0.173

NOTE: Bold numbers represent values higher than 0.2. The correlation coefficients that are significant at the 0.01 level are identified with two asterisks, and those significant at the 0.05 level are identified with one asterisk.

DISCUSSION AND CONCLUSIONS

Several authors in the area of activity-based modeling have argued that the generation of activities should be based on needs (*6, 7, 9*). However, systematic empirical research on exactly which needs are responsible for which activities individuals conduct in daily life had not been carried out. This paper makes a first attempt to elicit the needs underlying activity programming of social and leisure activities. Use of qualitative face-to-face interviews based on cognitive mapping techniques (*10–12*) and an Internet questionnaire to determine which needs are similar or have less influence revealed six needs that are distinct in their relationships with activities. The needs are social contact, physical exercise, relaxation, fresh air and being outdoors, new experiences, and entertainment. The activity profiles of these needs confirm the hypothesis that many-to-many relationships exist between activities and needs that give rise to substitution relationships between activities. In a third survey, statements were included and measurement scales were developed for every need. The scores on the needs scales were correlated with characteristics of the respondents and the activities they usually conduct every week. Analysis showed intuitive and interesting results. For instance, individuals having a higher need for social contact spend more time on social activities like visiting relatives or friends and going to a café, bar, or discothèque. The elderly tend to have fewer needs in general.

Important findings are that a substantial part of factors underlying activity choice can be interpreted as needs, the set of these needs is limited, and complex relationships between activities and needs exist and may give rise to negative generation effects between activities. The contribution of the paper to activity-based modeling is twofold. First, it represents the next step in developing a needs-based model of activity generation. This approach may find wider explanation in both cross-sectional activity-based models and in the further development of dynamic activity-based models. Second, it provides evidence of the relationship between activities and underlying needs, increasing the complexity of previous analyses and models of activity programming and duration choices. Future research will further analyze the data collected in the main questionnaire and estimate the parameters of a need-based model.

The study had some limitations. First, a replication of, in particular, the first survey is needed to verify completeness of the needs identified or whether additional dimensions play a role. Second, the present study focused on the nature of activities only, whereas other facets and, in particular, location may influence relationships between activities and needs as well. For example, the extent to which a given activity influences a need for experiencing fresh air and a green environment, obviously, will also depend on attributes of the location. Third, future research could focus on the extent to which travel mode and route choices for trips influence needs and, hence, interact with activities at

locations. For example, use of a bike may satisfy needs for physical exercise and being in open air and have a negative generation effect on recreation activities. Fourth, attributes of the residential location may have an influence on needs underlying activities, which had limited consideration in the present study. For example, attributes of residential location may satisfy or induce needs and, thus, have an impact on activity choices. Finally, reverse relationships may exist between activities and needs in that an activity can induce rather than satisfy a need. Such relationships are important because they give rise to positive generation effects between activities. To cover this aspect, an extension of the survey instrument is needed.

REFERENCES

1. Doherty, S. T. How Far in Advance Are Activities Planned? Measurement Challenges and Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1926,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 40–49.
2. Doherty, S. T., and A. Mohammadian. The Validity of Using Activity Type to Structure Tour-Based Scheduling Models. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
3. Habib, K. M. N., and E. J. Miller. Modelling Daily Activity Program Generation Considering Within-Day and Day-to-Day Dynamics in Activity-Travel Behaviour. *Transportation,* Vol. 35, No. 4, 2008, pp. 467–484.
4. Roorda, M. J., E. J. Miller, and K. M. N. Habib. Validation of TASHA: A 24-Hour Activity Scheduling Microsimulation Model. *Transportation Research Part A,* Vol. 42, No. 2, 2008, pp. 360–375.
5. Chapin, F. S. *Human Activity Patterns in the City.* John Wiley and Sons, New York, 1974.
6. Miller, E. An Integrated Framework for Modelling Short- and Long-Run Household Decision-Making. Presented at Progress in Activity-Based Analysis Conference, Maastricht, Netherlands, 2004.
7. Axhausen, K. W. (ed.). *Moving Through Nets: The Physical and Social Dimensions of Travel.* Elsevier, Oxford, United Kingdom, 2006.
8. Meister, K., M. Krick, and K. W. Axhausen. A GA-Based Household Scheduler. *Transportation,* Vol. 32, No. 5, 2005, pp. 473–494.
9. Arentze, T. A., and H. J. P. Timmermans. A Need-Based Model of Multi-Day, Multi-Person Activity Generation. *Transportation Research B,* Vol. 43, No. 2, 2009, pp. 251–265.
10. Reynolds, T. J., and J. Gutman. Laddering Theory, Method, Analysis, and Interpretation, *Journal of Advertising Research,* Vol. 28, No. 2, 1988, pp. 11–31.
11. Arentze, T. A., B. G. C. Dellaert, and H. J. P. Timmermans. Modeling and Measuring Individuals' Mental Representations of Complex Spatio-Temporal Decision Problems. *Environment and Behavior,* Vol. 40, No. 6, 2008, pp. 843–869.
12. Dellaert, B. G. C., T. A. Arentze, and H. J. P. Timmermans. Shopping Context and Consumers' Mental Representation of Complex Shopping Trip Decision Problems. *Journal of Retailing,* Vol. 84, No. 2, 2008, pp. 219–232.
13. Timmermans, H. J. P., T. A. Arentze, M. Dijst, E. Dugundji, C.-H. Joh, L. Kapoen, S. Krijgsman, K. Maat, and J. Veldhuisen. Amadeus: A Framework for Developing a Dynamic Multi-Agent, Multi-Period, Activity-Based Micro-Simulation Model of Travel Demand. Presented at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.
14. Sun, Z., T. A. Arentze, and H. J. P. Timmermans. A Heterogeneous Latent Class Model of Activity Rescheduling, Route Choice and Information Acquisition Decisions Under Multiple Uncertain Events. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
15. Adler, T., L. Rimmer, and D. Carpenter. Use of Internet-Based Household Travel Diary Survey Instrument. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1804,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 134–143.
16. Bricka, S., and J. Zmud. Impact of Internet Retrieval for Reducing Nonresponse in a Household Travel Survey. Presented at 82nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Factors Influencing the Planning of Social Activities

Empirical Analysis of Data from Social Interaction Diaries

Pauline van den Berg, Theo Arentze, and Harry Timmermans

Results of a study on the planning of social activities are reported. Data collected in the Netherlands from social interaction diaries were used to estimate a multinomial logistic regression model to analyze whether a social activity is prearranged, routine, or spontaneous as a function of personal and household characteristics, social activity characteristics, and characteristics of the contacted person. The results show that the planning of social activities is significantly influenced by gender, presence of children, education level, income, and time spent on work and school. Social activity characteristics were also found to have a significant effect. Social activities scheduled later in the day are less likely to be routine. In contrast, social activities of longer duration and taking place on the weekend are more likely to be routine or planned. The location, the main purpose of the social interaction, and detailed characteristics of with whom the social interaction took place were also found to affect the scheduling process significantly.

In the past decade, the research community in travel behavior has demonstrated an increasing interest in the planning or scheduling of activities (*1–8*). It is generally believed that a better understanding of activity scheduling processes will contribute to the development of activity scheduling models (*9–11*) and an understanding of short-term dynamics (*12*).

However, within that research, social activities have gained only little, and not very detailed, attention. Social activities are usually viewed as a single category or even merged with leisure activities, neglecting their heterogeneity (*13*). Moreover, very little is known in travel behavior research about the influence of the social context (with whom) on planning behavior for different (social) activities (*14, 15*). As social activities are always performed with other people, their planning and scheduling decisions are influenced not only by the characteristics of individuals but also by the characteristics of the persons with whom they interact.

Posing that more knowledge about social activities and commitments is required, this study aims at increasing understanding of the planning of different types of social activities. To that end, a multinomial logistic regression model is estimated to investigate how personal and household characteristics, social activity characteristics, and characteristics of the contacted persons influence the planning of social activities: Is the social activity prearranged, routine, or spontaneous?

The analysis is based on data collected in the Netherlands by using a social interaction diary. This unique data set allows examination of activity planning behavior for social interactions in more detail than existing studies had done.

Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands. Corresponding author: P. van den Berg, p.e.w.v.d.berg@tue.nl.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 63–70.
DOI: 10.3141/2157-08

LITERATURE REVIEW

It has long been recognized in transportation research that travel is induced by activities. Therefore, insight into individuals' daily activity patterns and their activity planning and scheduling can contribute to the understanding of travel demand. Activity scheduling is the process of how individuals organize their activities and travel (*9*). The decisions that are made in an activity scheduling process include which activities to perform, where, which transport mode (and route) to use, at what time, for what duration, and with whom (*2*). Activity planning concerns the extent to which activities are preplanned.

Activity Planning

Activity planning has received increasing attention in the field of travel behavior research in the past decade. Mohammadian and Doherty used a mixed logit model to estimate if activities were planned weeks or months ago, the same week, or the same day or were impulsive (*4*). They found that activities are more likely to be preplanned (long in advance) if they take place out of the home, they occur early in the day, the duration and travel time are longer, and the number of people involved is higher. Older people, cell phone users, and females were found to be more likely to preplan their activities. The results of Mohammadian and Doherty indicate that compared to other types of activities, social activities are less likely to be preplanned long in advance.

Lee and McNally examined the processes of household activity planning and trip chaining by using data collected with a computerized survey instrument (*5*). They estimated ordered logit models to identify the factors that influence the planning horizon of activities. They found that daily schedules often start with certain activities, such as work and social activities, occupying a portion of the schedule. Other activities are then (impulsively) arranged around these preplanned activities. Activities of shorter duration are more likely to be spontaneously inserted in a schedule. They found that persons with children often expect more constraining activities. Females were found to be more structured in planning their week. Activities

at more-distant locations are planned earlier than are activities at closer locations.

Van Bladel et al. estimated a mixed logit model to assess the personal, household, activity, and schedule attributes that affect the decision of whether to plan activities beforehand (*7*). Their results are similar to Mohammadian and Doherty's findings. Activities early in the day, longer activities, and activities with more people and with household members were found to be more likely to be planned beforehand. Overall, 76% of the activities were found to be preplanned. Social activities were among the most impulsively performed activity types. Individual and household characteristics were not found to significantly affect the probability of activity planning.

Kang et al. estimated a bivariate probit model to examine the planning priority of joint activities for female and male household heads, respectively (*8*). The explanatory variables included activity attributes and personal and household characteristics. They found that female household heads preplanned 75% of their joint activities and male household heads preplanned 71% of their joint activities. This finding is, like their other findings, highly consistent with Mohammadian and Doherty's (*4*). Besides the regular results (a positive effect for longer duration and travel time), they found significant effects for household composition, dwelling type, and urban versus suburban areas.

In those studies discussed, social activities usually are treated as one of the activity types. However, this paper argues that social activities are not homogeneous and should not be treated as a single category. There will be differences in planning behavior for different types of social activities. That social activities are not homogeneous is recognized by Kemperman et al. (*13*), who examined the differences in the scheduling of various social activities. Their results show that there are distinct scheduling patterns for different kinds of social activities, which underlines that social activities should not be treated as homogeneous.

Social Dimension in Activity Planning Behavior

The role of the social dimension has received increasing attention in travel behavior research (*14*). Social influence in travel-related decisions has been studied by Dugundji and Walker (*15*) and by Páez and Scott (*16*). The generation of social activities and the spatial distribution of social networks have been studied comprehensively by Carrasco and Miller (*17, 18*) and Carrasco et al. (*19*). Arentze and Timmermans proposed a theoretical framework to incorporate social activities and changing social networks in the microsimulation of activity-travel patterns (*20*).

Furthermore, the role of companion type (including nonhousehold members) in the participation and scheduling decisions regarding joint activities has received attention. For instance, Goulias and Kim studied the role of "with whom" in selecting different activity types (*21*). Srinivasan and Bhat studied the impacts of demographic and activity characteristics on the choice of companion type for leisure activities by using data from the American Time Use Survey (*22*). Kapur and Bhat used the same data to study the company type in weekend discretionary activity participation patterns (*23*). Habib et al. studied the relationship between the social context (with whom respondents interact) and the start time and duration of social activities (*24*). Analyzing activity diary data from Toronto, they showed that significant correlations exist between the social category of with whom social activities are performed and the start time and the duration of social activities. They further indicated that scheduling decisions for social activity are influenced more by with whom people socialize than by travel time or distance.

Despite an increasing interest in the influence of the social context on activity and travel behavior, little is known in travel behavior research about the influence of social context on the planning of different (social) activities. The social context is additionally important for the planning of social activities, as these activities are always performed with other people. On the one hand, the preceding section showed that the involvement of other people in an activity increases the odds of the activity being planned. On the other hand, social activities are found to be among the most impulsively performed activities. Either way, the literature review shows that more knowledge is needed about the relationship between characteristics of with whom and the planning of different social activities. This study will contribute to this knowledge by analyzing whether social activities are preplanned, routine, or spontaneous.

OBJECTIVES AND HYPOTHESES

Social activities are necessary for the maintenance of social networks. From an activity-based perspective, it can be assumed that after taking care of mandatory and maintenance activities, some time is left for people to become involved in various social activities. The objective for the present study is to analyze the effect of several characteristics on the odds of a social interaction being preplanned, routine, or spontaneous.

This classification was introduced by Gärling et al. (*25*), who argued that successfully forecasting travel behavior relies on a distinction among planned, habitual, and impulsive travel. This distinction has been used by other researchers (*26, 27*). To distinguish among these three classes, the first question is whether an intention is formed and the second, if it is formed, how much planning is involved (*25*). In the case of habitual (or routine) behavior, no intention is formed. In the case of impulsive behavior, an intention is formed; however, it is formed shortly before the activity and therefore very little planning takes place. Preplanned behavior is to a larger degree preceded by a deliberate decision. In general, the need for more-deliberate decisions or planning in advance is likely to be higher if the scheduling complexity and effort and the costs of the activity are higher.

How social activities are scheduled will vary by personal and household variables, such as age, gender, household composition, and social status. Although Van Bladel et al. did not find significant effects for individual and household characteristics (*7*), it is assumed here that there is a personal preference for planning social activities that can be explained by sociodemographic variables. Age and gender are two variables that may influence whether social activities are routine, prearranged, or spontaneous. Household composition may have an effect on the planning of social activities as well. The presence of a partner and, especially, children in the household may be a motivation to keep to a routine. Furthermore, the planning of social activities may require more effort if the schedules of family members have to be considered. The presence of a partner and children might thus decrease the odds of having spontaneous social contacts. The planning preference for social activities might also be affected by social status variables such as education level, income, work hours, and car ownership. It is assumed that people with more free time (less work or study) have more possibilities for spontaneous social interactions.

The second set of variables that can explain preference in social activity planning consists of the characteristics of the social activity. The start time and the duration of the social activity are factors that may significantly affect the planning behavior. It is hypothesized here that social activity duration has a positive effect on the social interactions being preplanned, because longer social activities require more planning effort. It is further expected that social interactions are more likely to be impulsive when people have more free time. Therefore, it is hypothesized that during the weekend, a lower preference for planning will be found. The location of the social interaction is also likely to affect social activity planning.

Furthermore, in addition to these more-general factors, this study attempts to explicitly analyze the role of several characteristics that are specific for social activities. This study acknowledges heterogeneity in social activities. The unique data set makes it possible to study the effect of the main purpose of the social activity on the planning behavior. The main purpose can be just talking or chatting, performing a joint activity, visiting or hosting, a short question or message or making an appointment, discussing or exchanging information or advice, or some other purpose. It is expected that social interactions for the purpose of a joint activity or a visit are more likely to be prearranged than would be a social interaction for the purpose of just chatting or talking.

As discussed earlier, social activities always take place with one or more other persons. Social activities with more persons will require more scheduling effort and will therefore be likely to be preplanned. The planning preference for social activities is likely to be affected by characteristics of the contacted person. For instance, the social-role relationship with the contacted person is likely to affect the planning behavior. Social activities with neighbors tend to require less planning than social activities with others. There will also be a higher chance of running into a neighbor and having a spontaneous chat compared to other social categories. Other characteristics of the alter that may be of influence are gender, age, how long ego and alter have known each other, the strength of their relationship, and their contact frequency. Finally, the planning preference for social activities is likely to increase if the distance between the homes of the ego and the alter increases, because of increasing travel costs.

DATA

The data used for this study were collected between January and June 2008 in a number of neighborhoods in and around Eindhoven, the Netherlands. The data collection instrument consisted of a 2-day paper-and-pencil social interaction diary. The social interaction diary consisted of three parts. In the first part, the respondents were asked to record all their social interactions during 2 days. Social interactions were defined as all forms of social contact, for instance, visiting, performing a joint activity, having a conversation (face-to-face as well as over the phone or online), sending or receiving an e-mail, an SMS, a letter, or a fax. Interactions at home with members of the household were not included, nor were interactions as a customer or work-related interactions. Along with whether the social interaction was routine, prearranged, or spontaneous, information was gathered about the communication mode used for the interaction, the day of the week, the time and duration of the interaction, the purpose of the interaction, and with whom and where the interaction took place. If travel was involved, the travel distance and transport mode were recorded as well.

In the second part of the interaction diary, the respondents were asked to fill out a page with questions about every person they interacted with during those 2 days. These questions included age, gender, and social category of the person, how long the respondent has known the person, the strength of the relationship with the person, the distance between the homes of the respondent and the person, and the frequency of social contact with the person by different communication modes.

The third part of the interaction diary contained a questionnaire on personal and household characteristics of the respondents (age, gender, household composition, level of education, income, occupation, sports and hobbies, involvement in clubs or unions, transportation modes, and time pressure), characteristics of the environment (urban density, facilities in residential location), and their access to and use of information and communication technologies (ICT) (computer, Internet, and mobile phone).

To ease respondent burden, the interface of the diary was kept as simple as possible. The respondents were asked to record their social interactions as soon as possible after they occurred. However, they also received an additional interaction worksheet that they could use during the day to remember their interactions, in case it was impossible to take the booklet along.

To increase respondent participation, people were asked in person if they were willing to participate in this study. If they were, they were given a social interaction diary that was collected approximately a week later. Of 3,699 people who answered their door, 1,648 (45%) accepted a diary. From these, 747 useful diaries were returned, an overall response rate of 20%. Details about the data collection procedure are discussed elsewhere (*28*).

The data were collected in a larger study of the link between properties of the built environment, ICT use, social networks, and social travel patterns. Unfortunately, the data set does not provide full diary data or detailed information on the social activity planning process, such as the time between planning and occurrence of a social interaction, or modification of the planned activities. It does, however, allow analysis of the factors that influence the odds of a social interaction being routine, prearranged, or spontaneous.

ANALYSIS AND RESULTS

The analysis focuses on predicting whether face-to-face social activities are routine, prearranged, or spontaneous based on a set of explanatory variables. From 747 respondents, 4,177 face-to-face social interactions were recorded (initiated either by the respondent or by a contact person). However, a number of cases contained missing values in the characteristics of with whom, especially the face-to-face contact frequency with this person. After deletion of cases with missing values, 2,505 useful cases were entered in the analysis.

A little less than half of these social interactions were prearranged. One-fifth were routine, and one-third were spontaneous. This observation can be seen in the following table:

	N	*Percentage*
Routine	500	20
Prearranged	1,143	46
Spontaneous	862	34
Total	2,505	100

The relatively high percentage of spontaneous social interactions is in line with findings that social activities are among the most impulsively performed activities (*4, 7*).

The objective for this study is to analyze the effects of personal and household characteristics, the social activity characteristics, and the characteristics of the contacted person on the odds of the social activity being prearranged or routine relative to being spontaneous. As the dependent variable in the model is a nominal variable, a multinomial logistic regression model is estimated to analyze these relationships. The last category, spontaneous, serves as the reference category in the model. This means that the estimation results for routine and prearranged should be interpreted relative to spontaneous.

To obtain a parsimonious model, backward elimination was used to remove the coefficients that were not statistically significant at the 0.1 significance level in the likelihood ratio test. The explanatory variables that are included in the model are shown in Table 1.

The results of the multinomial logistic regression model for social activity planning behavior can be seen in Table 2. The results of the model show a reasonable fit to the data with many significant coefficients.

For personal and household characteristics, the results indicate that males are more likely than females to have routine social interactions relative to spontaneous social interactions. The presence of children in the household is found to have a positive effect on social interactions being routine or preplanned. This suggests that caring for children makes it more difficult to carry out social interactions spontaneously. For higher education, there is a negative coefficient for routine and a positive coefficient for preplanned. However, the significance level for both coefficients is above 10%. People with a high income are found to be significantly less likely to have routine social interactions. The number of hours spent on studying is found to have a positive effect on social activities being routine. Although not significant, the results show that the amount of time spent on working on the diary day has a positive effect on routine and a negative effect on preplanned social activities.

The results indicate that almost all social activity characteristics significantly influence the planning preference. The negative coefficient for start time indicates that social activities that occur later in the day are less likely to be routine. Longer social activities were found to have a tendency to be routine or preplanned. This finding is in line with findings from other studies (*4, 7, 8*). Longer activities require more scheduling effort.

For the weekend, negative coefficients are found for routine and preplanned social interactions. This suggests that on the weekend, social activities are likely to be spontaneous. The same was found (for all activities) by Mohammadian and Doherty (*4*) and Van Bladel et al. (*7*). This is because people have more free time on the weekends and therefore have fewer fixed commitments and a higher possibility to perform impulsive activities.

For activity location, the results show negative coefficients for almost all locations relative to a home location. This indicates that social activities at home (of ego or alter) are most likely to be routine or prearranged. This appears to contrast with Mohammadian and Doherty's finding that out-of-home activities are less likely to be impulsive as they need more scheduling effort (*4*). However, for social interactions, it can be argued that spontaneous or coincidental interactions (running into someone) are more likely to happen out of the home. Note also that social interactions with only household members at home were not recorded in the social interaction diaries. The finding of a higher planning preference for social interactions at home is therefore convincing. The results show that social interactions at a place of entertainment are least likely to be routine. Social activities at work or at other locations are least likely to be prearranged.

For main purpose of the interaction, the results indicate that joint activities tend to be routine or prearranged. As expected, social interactions for the purpose of visiting tend to be prearranged. Social interactions for exchanging information or advice or discussing are more likely to be routine or preplanned, relative to social interactions for talking or chatting, which is an intuitive finding. Social interactions for other purposes have the highest positive coefficients for routine and prearranged, suggesting that they are least likely to be spontaneous or coincidental.

With whom a social interaction takes place is also found to influence whether the interaction was routine, preplanned, or spontaneous. The results show that social interactions with one person are less likely to be routine or prearranged compared to social interactions with two or more persons. This finding is consistent with findings by Mohammadian and Doherty (*4*) and Van Bladel et al. (*7*), who found that activities with more people are more likely to be planned beforehand.

The effect of a number of characteristics of the contacted person was tested as well for cases in which the social interaction took place with one person. Regarding social category of the contacted person, (almost) all the estimated coefficients are negative. This indicates that social interactions with a household member, a relative, a friend, a neighbor, or a colleague are less likely to be routine or prearranged, relative to an interaction with an acquaintance. The highest and most significant negative coefficients are found for a neighbor, which indicates that social interactions with a neighbor are most likely to be spontaneous.

Whereas the age of the respondent was not found to significantly influence planning behavior, the age of the alter was. If the alter is young (<30) the interaction tends to be prearranged. If the alter is between 30 and 49, the social activity is less likely to be routine. The results for strength of tie suggest that the stronger the tie between ego and alter, the more likely the social activity is routine or preplanned. However, this effect is to some extent corrected by the higher probability of routine social interactions with an alter that is not long known. The distance between the homes of the ego and the alter affects the planning preference as hypothesized. If the distance increases, the odds of a social activity being prearranged also increase. Finally, a high face-to-face contact frequency is found to have a positive effect on social activities being routine.

CONCLUSION AND DISCUSSION OF RESULTS

The planning and scheduling of activities has received increasing attention from travel behavior researchers in the past decade. However, the planning of social activities has received little attention. Moreover, social activities usually are treated as a homogeneous set of activities, negating the variety of these activities. Therefore, this paper sought to increase understanding of the planning preference for social interactions.

Data collected in the Netherlands from social interaction diaries were used to estimate a multinomial logistic regression model to analyze whether a social activity is prearranged, routine, or spontaneous. The explanatory variables in the model are personal and household characteristics, characteristics of the social activity, and characteristics of the contacted person. All three sets of explanatory variables were found to significantly influence the planning of social activities.

TABLE 1 Sample Characteristics

Characteristic	*N*	%	Mean	SD
Personal and Household Characteristics (*N* = 747 respondents)				
Male	293	39		
Child(ren) under 18 years in household	333	45		
Higher education (BSc or higher)	442	46		
High income (>€3,000 per month)	247	33		
Work that day (hours)			2.62	3.52
School that day (hours)			0.38	1.28
Social Activity Characteristics (*N* = 2,505 face-to-face social activities)				
Start time			13.83	3.99
Duration (min.)			79.06	106.39
Weekend	550	22		
Weekday	1,963	78		
Location				
Home of ego or alter	1,079	43		
Work	415	17		
Place of entertainment	129	5		
Sports	178	7		
Public space outside or on the road	262	10		
Other	450	18		
Main purpose				
Talk or chat	781	31		
Joint activity	521	21		
Visit	385	15		
Short question, message, or app.	191	8		
Information, advice, or discuss	286	11		
Other	349	14		
Characteristics of with Whom (*N* = 2,505 face-to-face social activities)				
With two or more alters	1,059	42		
With one alter	1,454	58		
Characteristics of with Whom in Case of One Alter				
Category				
One alter + household member	79	3		
One alter + relative	292	12		
One alter + friend	219	9		
One alter + neighbor	197	8		
One alter + colleague	279	11		
One alter + acquaintance or stranger	388	15		
Age				
One alter + young (<30)	296	12		
One alter + middle age (30–49)	677	27		
One alter + old (≥50)	481	19		
Tie strength				
One alter + very strong tie	363	14		
One alter + reasonably strong	451	18		
One alter + not strong	640	26		
Years known				
One alter + known <2 years	355	14		
One alter + known 2–5 years	322	13		
One alter + known 5–15 years	329	13		
One alter + known >15 years	448	18		
Distance: One alter + log (distance + 1)			0.72	0.58
Contact freq.				
One alter + >weekly face-to-face contact	636	25		
One alter + once a week face-to-face contact	298	12		
One alter + 1–3 per month face-to-face contact	277	11		
One alter + ≤once a month face-to-face contact	243	10		

TABLE 2 Multinomial Logistic Regression Model Estimates for Planning of Social Activities

	Routine		Prearranged	
	B	Sig.	B	Sig.
Constant	−0.912	0.007	−1.082	0.000
Personal and Household Characteristics				
Male	0.418	0.002	0.080	0.504
Child(ren) under 18 years in household	0.421	0.002	0.278	0.019
Higher education (BSc or higher)	−0.164	0.227	0.142	0.238
High income(>€3,000 per month)	−0.339	0.012	0.104	0.386
Work that day (hours)	0.039	0.104	−0.019	0.365
School that day (hours)	0.138	0.004	0.046	0.350
Social Activity Characteristics				
Start time	−0.041	0.019	−0.004	0.799
Duration (min.)	0.009	0.000	0.011	0.000
Weekend	−0.519	0.006	−0.494	0.001
Location				
At home of ego or alter (reference)				
Work	0.334	0.173	−1.077	0.000
Place of entertainment	−0.926	0.015	−0.670	0.017
Sports	−0.003	0.991	−0.525	0.028
Public space outside or road	−0.163	0.493	−0.821	0.000
Other	−0.281	0.167	−1.079	0.000
Purpose				
Talk or chat (ref.)				
Joint activity	1.117	0.000	1.959	0.000
Visit or host	−0.086	0.722	1.219	0.000
Short question, message, or app.	−0.235	0.336	0.346	0.124
Information, advice, or discuss	0.703	0.001	1.920	0.000
Other	1.786	0.000	2.650	0.000
Characteristics of with Whom				
With one alter	−1.609	0.000	−0.523	0.098
Characteristics of with Whom in Case of One Alter				
Category				
One household member	−0.848	0.128	−0.010	0.984
One relative	−0.714	0.074	−0.568	0.074
One friend	−1.219	0.001	−0.149	0.589
One neighbor	−1.527	0.000	−0.916	0.001
One colleague	−0.810	0.009	−0.682	0.022
One acquaintance or stranger (ref.)				
Age				
One alter + young (<30)	−0.388	0.153	0.533	0.023
One alter + middle age (30–49)	−0.399	0.055	0.129	0.475
Tie strength				
One alter + very strong tie	1.271	0.000	1.061	0.001
One alter + reasonably strong	0.411	0.070	0.372	0.075
Years known				
One alter + known <2 years	0.698	0.037	−0.163	0.550
One alter + known 2–5 years	0.343	0.274	−0.369	0.139
One alter + known 5–15 years	0.584	0.047	−0.312	0.181
One alter + known >15 years (ref.)				
Distance: log (distance + 1) between homes	0.083	0.680	0.492	0.003
Face-to-face contact frequency				
One alter + >weekly face-to-face contact	1.679	0.000	−0.072	0.783
One alter + once a week face-to-face contact	1.606	0.000	0.259	0.299
One alter + 1–3 per month face-to-face contact	0.664	0.106	0.160	0.512
One alter + ≤once a month face-to-face contact				

NOTE: *N*: 2,505 face-to-face social interactions, reported by 747 respondents; −2 log likelihood—intercept only: 5.243E3; final: 3.994E3; likelihood ratio tests—chi-square: 1.249E3; df: 72; sig.: .000; *R*-square—Cox & Snell: .393; Nagelkerke: .448; McFadden: .238.

The results indicate that preference in social activity planning is significantly influenced by gender, presence of children, education level, and income. The social activity characteristics were also found to affect the planning preference for social activities. Social activities that occur later in the day are less likely to be routine. Longer social activities and social interactions taking place on the weekend tend to be routine or preplanned. Social activities at a home location tend to be routine or prearranged, whereas social activities at work or at other locations are least likely to be prearranged. Social interactions at a place of entertainment are least likely to be routine.

The detailed characteristics of social interaction that make this study unique (the main purpose of the social interaction and characteristics of with whom the social interaction took place) were found to significantly affect the planning preference. Joint activities are likely to be routine or prearranged. Social interactions for the purpose of visiting tend to be prearranged. Social interactions for exchanging information or advice or discussing and social interactions for other purposes are more likely to be routine or preplanned.

Social interactions with two or more persons are less likely to be spontaneous. The highest and most significant negative coefficients are found for social interactions with a neighbor. If the alter is young (<30) the interaction tends to be prearranged. If the alter is between 30 and 49, the social activity is less likely to be routine. The stronger the tie between ego and alter, the more likely the social activity is routine or preplanned. However, if the alter is not (or not long) known, there is a higher probability of the social interaction being routine or preplanned. If the distance between the homes of ego and alter increases, the odds of a social activity being prearranged increase.

In general, the findings are consistent with the hypotheses. It was hypothesized that the odds of preplanning would be higher if the scheduling complexity and effort and the costs of the activity are higher. This result is corroborated by the finding that social activities with longer durations, with more than one other person, and with people who live further away tend to be preplanned. It was further assumed that social interactions are more likely to be impulsive when people have more free time. This assumption was confirmed by the finding that social activities on the weekend tend to be spontaneous. The number of hours spent on studying and working were found to have a positive effect on social activities being routine.

Although the findings presented in this paper are interesting, the study has some limitations. The data on which the analysis is based were collected by using a social interaction diary. The data therefore provide limited information on the other activities performed during the diary days. Moreover, no strict definitions of routine, prearranged, and spontaneous were given to the respondents. Therefore, respondents may have interpreted these classes differently (e.g., regarding the person with whom the interaction took place, the time of the interaction, or the location of the interaction). However, in most cases the planning process will capture all these aspects (with whom, the location, and the time) at the same time. Another limitation of this study is that detailed information on the planning process, such as the time between the planning and the occurrence of the social interaction, or modification of the planned activities, is not available in the data set.

Nevertheless, by using unique data with a specific focus on social activities, this study adds to the understanding of activity planning in relation to travel demand for social purposes, which has received little attention in academic research. This study indicates that activity-based models can be improved for social activities by a further segmentation of social activities and by giving more attention to the social context of these activities.

REFERENCES

1. Doherty, S., and E. Miller. A Computerized Household Activity Scheduling Survey. *Transportation,* Vol. 27, No. 1, 2000, pp. 75–97.
2. Doherty, S., E. Miller, K. Axhausen, and T. Gärling. A Conceptual Model of the Weekly Household Activity-Travel Scheduling Process. In *Travel Behaviour, Spatial Patterns, Congestion, and Modelling* (E. Stern, I. Salomon, and P. Bovy, eds.), Edwards Elgar, Cheltenham, United Kingdom, 2002, pp. 233–264.
3. Lee, M., and M. McNally. On the Structure of Weekly Activity/Travel Patterns. *Transportation Research Part A,* Vol. 37, No. 10, 2003, pp. 823–839.
4. Mohammadian, A., and S. T. Doherty. Mixed Logit Model of Activity-Scheduling Time Horizon Incorporating Spatial–Temporal Flexibility Variables. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1926,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 33–40.
5. Lee, M., and M. McNally. An Empirical Investigation on the Dynamic Processes of Activity Scheduling and Trip Chaining. *Transportation,* Vol. 33, No. 6, 2006, pp. 553–565.
6. Ruiz, T., and M. J. Roorda. Analysis of Planning Decisions During the Activity-Scheduling Process. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 46–55.
7. van Bladel, K., T. Bellemans, D. Janssens, and G. Wets. Activity Travel Planning and Rescheduling Behavior: Empirical Analysis of Influencing Factors. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2134,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 135–142.
8. Kang, H., D. M. Scott, and S. T. Doherty. Investigation of Planning Priority of Joint Activities in Household Activity-Scheduling Process. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2134,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 82–88.
9. Gärling, T., T. Kalén, and J. Romanus. Computer Simulation of Household Activity Scheduling. *Environment and Planning A,* Vol. 30, No. 4, 1998, pp. 665–679.
10. Arentze, T., and H. Timmermans. *Albatross: A Learning-Based Transportation Oriented Simulation System.* European Institute of Retailing and Services Studies, Eindhoven, Netherlands, 2000.
11. Miller, E. J., and M. J. Roorda. Prototype Model of Household Activity-Travel Scheduling. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1831,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 114–121.
12. Joh, C.-H., T. Arentze, and H. Timmermans. Activity-Travel Scheduling and Rescheduling Decision Processes: Empirical Estimation of Aurora Model. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1898,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 10–18.
13. Kemperman, A., T. Arentze, and H. J. P. Timmermans. Social Commitments and Activity-Travel Scheduling Decisions. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1977,* Transporation Research Board of the National Academies, Washington, D.C., 2006, pp. 242–249.
14. Axhausen, K. Social Networks and Travel: Some Hypotheses. In *Social Aspects of Sustainable Transport: Transatlantic Perspective* (K. Donaghy, ed.), Ashgate, Aldershot, United Kingdom, 2002, pp. 98–110.
15. Dugundji, E., and J. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1921,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.
16. Páez, A., and D. Scott. Social Influence on Travel Behavior: A Simulation Example of the Decision to Telecommute. *Environment and Planning A,* Vol. 39, No. 3, 2007, pp. 647–665.
17. Carrasco, J., and E. Miller. The Social Dimension in Action. A Multilevel, Personal Networks Model of Social Activity Frequency Between

Individuals. *Transportation Research Part A,* Vol. 43, No. 9, 2009, pp. 90–104.
18. Carrasco, J., and E. Miller. Exploring the Propensity to Perform Social Activities: A Social Networks Approach. *Transportation,* Vol. 33, No. 5, 2006, pp. 463–480.
19. Carrasco, J. A., E. J. Miller, and B. Wellman. How Far and with Whom Do People Socialize? Empirical Evidence About the Distance Between Social Network Members. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2076,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 114–122.
20. Arentze, T., and H. Timmermans. Social Networks, Social Interactions, and Activity-Travel Behavior: A Framework for Microsimulation. *Environment and Planning B,* Vol. 35, No. 6, 2008, pp. 1012–1027.
21. Goulias, K. G., and T.-G. Kim. Analysis of Activity Type Classification and Issues Related to With-Whom and For-Whom Questions of an Activity Diary. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2005.
22. Srinivasan, S., and C. Bhat. Companionship for Leisure Activities: An Empirical Analysis Using the American Time Use Survey. *Proc., Innovations in Travel Demand Modeling Conference,* Austin, Tex., 2006. http://onlinepubs.trb.org/onlinepubs/conf/CP42v2.pdf.
23. Kapur, A., and C. R. Bhat. Modeling Adults' Weekend Day-Time Use by Activity Purpose and Accompaniment Arrangement. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2021,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 18–27.
24. Habib, K., J. M. N. Carrasco, and E. A. J. Miller. Social Context of Activity-Scheduling: Discrete–Continuous Model of Relationship Between "with Whom" and Episode Start Time and Duration. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2076,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 81–87.
25. Gärling, T., R. Gillholm, and A. Gärling. Reintroducing Attitude Theory in Travel Behavior Research: the Validity of an Interactive Interview Procedure to Predict Car Use. *Transportation,* Vol. 25, No. 2, 1998, pp. 129–146.
26. Jakobsson, C. Accuracy of Household Planning of Car Use: Comparing Prospective to Actual Car Logs. *Transportation Research Part F,* Vol. 7, No. 1, 2003, pp. 31–42.
27. Lee-Gosselin, M., and L. Miranda-Moreno. What Is Different About Urban Activities of Those with Access to ICTs? Some Early Evidence from Québec, Canada. *Journal of Transport Geography,* Vol. 17, No. 2, 2009, pp. 104–114.
28. Van den Berg, P., T. Arentze, and H. Timmermans. Social Networks, ICT Use and Activity-Travel Patterns. Data Collection and First Analyses. Presented at 9th International Conference on Design and Decision Support Systems in Architecture and Urban Planning, Netherlands, 2008.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Genetic Algorithm to Estimate Cumulative Prospect Theory Parameters for Selection of High-Occupancy-Vehicle Lane

Joseph Y. J. Chow, Gunwoo Lee, and Inchul Yang

Recent literature suggests a need for a more realistic representation of driver behavior. In an effort to integrate prospect theory as a potential descriptive method into traveler behavior, it is important to determine the validity of previous estimated parameters obtained from empirical studies and interviews of people who are not in the same fast-paced, dynamic, travel time disutility setting as real drivers. A genetic algorithm is used to simultaneously estimate the parameters of the cumulated prospect theory (CPT) value and weight functions as well as the coefficients of the random utility model; this procedure leads to estimates that have a higher likelihood value and statistical significance than an equivalent expected utility–based logit model or a CPT-based logit model using the empirical values developed earlier. The value function parameters generally conform to conclusions from previous literature. The weight function parameters, however, suggest that drivers in a fast-paced changing environment with multiple subjects for prospect evaluation may become overwhelmed by the certainty effect.

Strictly rational traveler behavior has been the foundation of many studies conducted in the field of transportation. However, the works of Simon (*1*) and Kahneman and Tversky (*2*) proved that human behavior is not strictly rational and exhibits many elements that are driven by uncertainty and risk averseness. This descriptive behavior, as opposed to the normative behavior associated with strict rationality, is a more realistic representation of how people behave.

Numerous researchers into travel behavior have explored the many facets of these concepts. Mahmassani and Chang applied Simon's concept of bounded rationality to Wardrop's strictly rational user equilibrium (*3*). Bonsall (*4*) and de Palma and Picard (*5*) verified the risk-seeking attitude of drivers in situations of disutility. Chen et al. looked at the interaction of drivers with different classes of risk attitudes in assessing travel time reliability (*6*). Ben-Akiva et al. extended the discrete choice model to include descriptive behavior by using latent variables (*7*). Hess et al. explored the asymmetric nature of preferences in estimating willingness-to-pay with discrete choice models (*8*).

In practice, the argument is one of simplicity versus realism. Proponents of model simplicity would argue that with more-complex transport problems—multiple criteria such as sustainability and social justice, real-time traveler information—the added realism of descriptive behavior might not be worth the additional model complexity. However, other types of transport problems rely on a realistic traveler behavior, such as those with high sensitivity to traveler choices (e.g., dynamic toll pricing) or those with high consequences that need to capture different risk attitudes (e.g., disaster evacuation).

Institute of Transportation Studies, University of California, Irvine, 4000 Anteater Instruction and Research Building, Irvine, CA 92697-3600. Corresponding author: J. Y. J. Chow, ychow@uci.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 71–77.
DOI: 10.3141/2157-09

Recent research has sought to integrate descriptive theories such as prospect theory (PT) into traveler choice models (*2*). In addition to willingness-to-pay (*8*), wait time (*9*), travel time reliability (*10*), departure time (*11*), and route choice (*12, 13*) have been considered by other recent studies. Many studies assume the same five parameter values estimated empirically by Kahneman and Tversky (*2*).

However, it has been pointed out (*11–13*) that the parameters can differ depending on the choice of reference point in evaluating utility or disutility, the setting of the choices being made such as peak or off-peak travel, and other constraints that exist for the utility (e.g., early arrival is not always desired). Senbil and Kitamura tried to overcome this problem by estimating the parameters for one of the two functions that define PT (*11*). They estimated only the parameters of the value function while leaving the weight function with the same empirical values.

This study takes the estimation a step further by using a genetic algorithm (GA) to simultaneously estimate the parameters for both functions. An experiment is constructed to illustrate how to estimate a discrete choice model for choice of high-occupancy-vehicle (HOV) lane based on current and perceived speeds. The likelihood function is compared between a simple binary logit model based on traditional expected utility, one with the nonlinear representative utility based on PT with the empirical values and the other with PT estimated with the GA.

PROSPECT THEORY VERSUS EXPECTED UTILITY THEORY

Expected utility (EU) theory is based on evaluating the utility from a set of outcomes as the expected value of the outcomes. For example, consider a situation in which a driver has the option of traveling on the HOV lane or in the mixed-use lane because he has an additional passenger on board. Say the driver can choose from the following options:

Option A. Stay on the mixed-use lane at a speed of 60 mph.

Option B. Go onto the HOV lane for a speed of 65 mph with 60% chance of slowing to 55 mph.

The expected utility of Option A can be formulated as 60, whereas the expected utility of Option B would then be 59. On the basis of expected utilities alone, the driver would rationally choose Option A because 60 > 59 and he is assumed to be maximizing his utility. However, what happens when more about the driver's situation is learned? Expected utility theory does not offer additional insight from information about the current speed of the driver. The theory has shown that human behavior would tend toward risk-seeking attitudes in situations of disutility, and it would not be unusual for a majority of drivers choosing Option B to view speed as a direct corollary to a disutility such as travel time.

Several effects explored by Kahneman and Tversky (*2*) under PT can account for such differences in behavior that EU cannot:

- Certainty effect. People place more weight on outcomes that are more certain (choose gain of \$100 versus probability of gaining \$110 with 95% chance and \$0 otherwise).
- Reflection effect. People are risk-seeking when prospects involve losses, as opposed to being risk averse when prospects involve gains.
- Isolation effect. Prospects can be decomposed into common and distinctive components, and different decompositions can lead to different preferences (common components can be discarded).

PT uses a two-phase approach that involves setting up an initial framing (editing) of the prospects to isolate the relevant components, and an evaluation phase that determines the preferred prospect. The first phase involves setting a reference point so that the relative gains and losses are used instead of absolute values, and the second phase involves interpreting the likelihoods of those relative values such that they realistically represent the asymmetries caused by the effects discussed earlier.

This method of decision making based on a reference point is shown by de Borger and Fosgerau to be a strong factor in explaining responses involving loss aversion in time and money (*14*). Their experiments show that drivers are more loss averse in the time dimension than in the cost dimension.

Cumulative prospect theory (CPT) is an updated version of PT that incorporates continuous distributions and any number of prospects. It extends the original version by allowing different decision weights for gains and losses. A detailed description of the method is provided by Tversky and Kahneman (*15*), although a brief summary is shown here to lead up to the experiment.

Let S be a finite set of states of nature and let the subsets be called events. Let X be a set of consequences or outcomes, which includes a neutral outcome denoted 0. All other outcomes are interpreted as gains or losses, denoted by positive or negative numbers, respectively. An uncertain prospect f is a function from S into X that assigns to each state $s \in S$ a consequence $f(s) = x$ in X.

To define a cumulative functional, the outcomes are ranked in increasing order. Define w as a function that assigns each $A \subset S$ a number $w(A)$ satisfying $w(\phi) = 0$, $w(S) = 1$, and $w(A) \geq w(B)$ whenever $A \supset B$. If a prospect f is given by a probability distribution $p(A_i) = p_i$, the A_i can be replaced by p_i. There would exist a strictly increasing value function v: $X \rightarrow \mathrm{Re}$, satisfying $v(x_0) = v(0) = 0$, and functions w^+ and w^- such that for $f = (x_i, p_i)$, $-m \leq i \leq n$,

$$V(f) = V(f^+) + V(f^-) \tag{1a}$$

$$V(f^+) = \sum_{i=0}^{n} \pi_i^+ v(x_i) \tag{1b}$$

$$V(f^-) = \sum_{i=-m}^{0} \pi_i^- v(x_i) \tag{1c}$$

where the decision weights $\pi^+(f^+) = (\pi_0^+, \ldots, \pi_n^+)$ and $\pi^-(f^-) = (\pi_{-m}^-, \ldots, \pi_0^-)$ for probabilistic prospects are defined by

$$\pi_n^+ = w^+(p_n) \tag{2a}$$

$$\pi_{-m}^- = w^-(p_{-m}) \tag{2b}$$

$$\pi_i^+ = w^+(p_i + \cdots + p_n) - w^+(p_{i+1} + \cdots + p_n) \quad 0 \leq i \leq n-1 \tag{2c}$$

$$\pi_i^- = w^-(p_{-m} + \cdots + p_i) - w^-(p_{-m} + \cdots + p_{i-1}) \quad 1-m \leq i \leq 0 \tag{2d}$$

Given preference homogeneity, v can be represented by a two-part power function:

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\beta & \text{if } x < 0 \end{cases} \tag{3}$$

Tversky and Kahneman (*15*) also fitted a curve to the w^+ and w^- functions:

$$w^+(p) = \frac{p^\gamma}{\left(p^\gamma + (1-p)^\gamma\right)^{1/\gamma}} \tag{4a}$$

$$w^-(p) = \frac{p^\delta}{\left(p^\delta + (1-p)^\delta\right)^{1/\delta}} \tag{4b}$$

The five parameters mentioned earlier are in Equations 3 and 4. The α and β parameters in the value function represent sensitivity to gains and losses, respectively. A value of less than 1 indicates diminishing sensitivity to the value. The λ parameter represents the loss aversion, with a higher value indicating greater aversion to loss. The γ and δ parameters represent the nonlinear fitting of weight functions for the probability distributions for gains and losses. The median values estimated empirically from the surveys by Kahneman and Tversky (*2, 15*) are $\alpha = \beta = 0.88$, $\lambda = 2.25$, $\delta = 0.69$, and $\gamma = 0.61$.

EXPERIMENT

HOV Lane Choice

One of the current problems in designing HOV lanes is that most models used to capture a user's choice of the mode (mixed-use lane or HOV lane) are dependent on aggregate trip-end demographics, such as median income, age, or gender. These discrete mode choice models (*16–18*) designate a user to an HOV lane for the duration of a trip without any consideration of traffic flow conditions.

These trip-end models cannot predict a user's mode of travel or the proportion of HOV users given a specific freeway segment with certain flow characteristics, such as during lane-specific incidents. With the rise of new HOV and high-occupancy toll policies, it is important to have more-sophisticated driver behavioral models.

The benefit of a performance-based mode choice model with lane-changing characteristics is that total vehicle miles traveled can be more realistically estimated, and the impact of incidents on HOV lane choice can be determined. Two different freeway segments that share the same characteristics in every way except incident occurrence should exhibit different rates of HOV usage. These benefits would be more apparent on freeways where traffic conditions change rapidly and many lane entry and exit points are available, such as areas with diverse land uses.

Choudhury showed that previous lane-changing models are significantly affected by exclusive lanes and developed an expected utility model that explicitly considers HOV lanes with latent variables (*19*). The model is composed of a lane-selection phase and a gap-acceptance phase. Although latent variables are used to account for heterogeneity in the unobservable lane selection decision making, it maintains the use of the normative EU.

Experiment Description

In this experiment, the HOV lane choice of observed traffic on a segment of I-80 is estimated for a specific period. The sample is obtained from the FHWA Next Generation Simulation data set. The I-80 study area is in Emeryville, California, and the data set was collected in April 2005 by FHWA.

The data were recorded by seven video cameras installed on a 30-story building, Pacific Park Plaza, at 6363 Christie Avenue. The I-80 trajectory data set has three data sets of 15-min intervals (4:00 to 4:15 p.m., 5:00 to 5:15 p.m., and 5:15 to 5:30 p.m.), for 45 min total. The study site is approximately 1,650 ft long and has six mainlines, including an HOV lane and an on-ramp at Powell Street (*20*).

The HOV lane in I-80 has continuous access, which means that sufficiently occupied vehicles can move into the HOV lane at any point. In other words, there are no specific access points for HOV lane access, so assumptions need to be made about drivers' average travel speed as well as their perception of upcoming driving speeds on the shared lanes and on the HOV lane given a highway segment such as the one being studied.

In this experiment, it is assumed that the travel speed and perceived travel speed distributions determined when the vehicle is located at one-quarter of the length of the segment are representative of their choice for that segment. The average historical speed of the traveler from the entry of the segment to the one-quarter length is compared to the distribution of downstream speeds on the shared lanes versus the distribution of speeds on the HOV lane. Instantaneous speeds of vehicles in front of the driver are used to construct a histogram of 5-mph intervals up to 45 mph, with a conditional expected speed value for the interval of speeds greater than 45 mph for each alternative.

For the model estimations, two data sets (5:00 to 5:15 p.m. and 5:15 to 5:30 p.m.) are utilized. The first 15-min data set (4:00 to 4:15 p.m.) is utilized for the models' validation, and the second period is used for parameter estimation. The statistics of the two time intervals' observed data sets are summarized as follows:

	Period	
	4:00 to 4:15 p.m.	*5:00 to 5:30 p.m.*
Switched to HOV lane	14	39
Stayed in mixed lanes	330	529
Total	344	568

A logit model was constructed by using the average speeds ahead of a driver on the shared lane and the HOV lane as the only independent variable. The representative utility was modeled by using CPT instead of EU. The utility function is shown in Equation 5 for both forms of tested logit models.

EU-based logit model:

$$U_{in} = \mu_1 + \mu_2 x_{in} + \epsilon_{in} \tag{5a}$$

CPT-based logit model:

$$U_{in} = \mu_1 + \mu_2 V_{in}\left(x_{in}, S_n\right) + \epsilon_{in} \tag{5b}$$

where

S_n = speed of the nth person in the N-sample data set,
x_{in} = speeds of the ith alternative for the nth person, and
μ_i = coefficients normalized as the differences between the alternative specific coefficients.

$V_{in}(x_{in}, S_n)$ can be obtained from Equation 1, which is broken down into the components in Equations 2 through 4. The *p*s in Equations 2 and 4 represent the distribution of speeds and can be represented by a 5-mph-interval histogram drawn from vehicles in front observed by the driver at the time of lane selection. ϵ_n is assumed to be a standard Gumbel distributed error term. The probability of person n choosing alternative i in a binary choice model as shown in (*21*) is:

$$P_n(i) = \Pr\left(U_{in} \geq U_{jn}\right) \tag{6}$$

The two alternatives in the binary choice model are mutually exclusive and exhaustive. Because of the Gumbel distributed error term, the probability of a driver choosing an HOV lane is shown in Equation 7, where u_{in} is the representative utility of alternative i for the nth person:

$$P_n(i) = \frac{1}{1 + e^{(u_{jn} - u_{in})}} \tag{7}$$

The current speed of the driver is used as the reference point for the PT. In the EU-based logit model, the V_{in}'s are simply the speeds x_{in} being perceived by the driver. The assumption is that drivers are fully aware of the probability distribution of downstream speeds when making their choices and therefore have no perception error. Going from a linear EU logit model to a CPT-based model, each $\mu_k x_{in}$ in Equation 5*a* is now replaced by a nonlinear representative utility function $\mu_k\, V(x_{in}, \alpha, \beta, \gamma, \delta, \lambda)$ as per Equations 1 through 4.

The parameters of the model are estimated by using maximum likelihood. As pointed out by Senbil and Kitamura (*11*), simultaneously estimating both the value and the weight functions for prospect theory by using derivative-based methods such as the Newton–Raphson method can lead to estimation problems. To avoid this, a GA is used as a heuristic to search for the optimal parameters by assuming that the probability distributions of value disturbances can be estimated with independent and identically distributed Gumbel variables. The results can be compared to the two control models: a logit model using EU and one using the empirical values from Tversky and Kahneman (*15*). If a heuristic can obtain a higher likelihood solution than the two control models, then there clearly exists a better choice of parameter values under the experimental setting.

The GA is chosen over other heuristics for convenience, although any heuristic that can handle nonconvex problems with continuous variables such as simulated annealing, tabu search, or radial basis functions may also be considered. A brief discussion of the GA is presented for context.

Genetic Algorithm

The maximum likelihood objective function involves estimating seven parameters from n samples. The solution requires taking the partial derivative of the likelihood function with respect to each parameter to satisfy first-order conditions. However, because of parameters such as γ and δ in Equation 6, one option is to apply the GA to obtain a maximum likelihood for the estimation of the CPT parameters.

Genetic algorithms, also known as global search algorithms, are heuristic search procedures based on the mechanics of natural selection and natural genetics, which include a genetic representation of solutions, an evaluation function for rating solutions, genetic operators to create offspring during reproduction, and values for the parameters of GAs.

Solutions can be represented by several encoding methods, such as binary encoding, real-number encoding, integer or literal permutation encoding, and general data structure encoding. The choice of encoding methods depends on the problem at hand. Genetic operators such as selection, crossover, and mutation play an essential role in searching the solution space both locally and randomly. In GAs, accumulated information is exploited by the selection mechanism, and new regions of the search space are explored by means of crossover and mutation (*22*).

In this experiment, real-number representation is adopted instead of binary representation so that each chromosome is encoded as a vector of real numbers. Real-number encoding is best used for function optimization problems. It has been widely confirmed that real-number encoding performs better than binary or Gray encoding for function optimization (*23*). Among various techniques of genetic operators and adaptation methods that have been proposed, examined, and compared since GAs were first introduced by Holland (*24*), Roulette wheel selection was chosen here as the selection method, convex crossover was chosen as the crossover method, and the conventional mutation method is used.

In roulette wheel selection, the chromosomes with high fitness values have more opportunities to be selected, so that the whole population tends to converge toward them, and finally the algorithms end up with premature consequences. To palliate the problem, some adaptive methods like transformation of the fitness values have been developed, such as windowing, exponential, linear transformation, and linear normalization. Detailed discussions of the transformation methods are available elsewhere (*25*). In this experiment, the linear normalization method is selected because of its good performance, shown in several test results.

The convex crossover is one of the arithmetical crossover methods whose basic concept is borrowed from convex set theory. In convex combination, the weighted average of two vectors is calculated as follows:

$$\lambda_1 x_1 + \lambda_2 x_2$$

where

$$\lambda_1 + \lambda_2 = 1 \qquad \lambda_1 > 0 \qquad \lambda_2 > 0 \tag{8}$$

TABLE 1 Estimated Parameters

Parameter	Range	Estimated Values via GA	Empirical CPT Values
α	[0,9]	0.7960	0.88
β	[0,9]	0.7848	0.88
γ	[0,9]	0.0147	0.61
λ	[0,9]	2.394	2.25
δ	[0,9]	5.479	0.69
ASC(μ_1)	[−9,9]	−2.9560 (8.105)	−2.7559
μ_2	[−9,9]	0.0466 (1.117)	0.0099

NOTE: () = *t*-statistics. ASC = alternative specific constant.

In convex crossover, the offspring created from two parents has a gene produced by the convex combination method with randomly generated weight factors.

RESULTS AND DISCUSSION

Because of GA, the maximum value of the log likelihood objective function obtained is −141.48 with the set of parameters as shown in Table 1. The range shown is the feasible range of values for the random generation of the parameter samples in the GA.

The empirical values of Kahneman and Tversky are shown in the right column for comparison (*2, 15*). The observed data are also estimated with the EU-based logit model and the CPT logit model with the fixed empirical values. The parameters for the EU and fixed CPT logit models are estimated by using Gauss, a statistical estimation software (*26*), and the results are shown in Tables 2 and 3, respectively. The *t*-statistics for the two coefficients (μ_1 and μ_2) in Table 1 are shown in parentheses. They were obtained by using Gauss with the five estimated parameters to estimate the logit function coefficients so that statistical significance can be determined. Compared to the likelihood values from EU (−142.07) and fixed CPT (−142.04), the estimated CPT value obtained with a heuristic results in a higher likelihood value of −141.48. This means that the true maximum likelihood is at least as high as that value. Even more significant, the *t*-statistics for the coefficients of alternative specific constant (μ_1) and the prospect value of speed (μ_2) are 8.105 and 1.117, respectively, which are much higher than for either the EU-based model (5.328 and 0.194) or the fixed CPT-based model (5.604 and 0.323).

A model validation is performed by using 344 sample data from the 4:15-to-4:30 p.m. period, which is assumed to behave similarly enough to the 5:00-to-5:30 p.m. period. The results of the validation

TABLE 2 Estimated Results of EU-Based Logit Model

Variable	Parameters	*t*-Statistics
HOV alternative specific constant (μ_1)	−2.700	5.328
Speed difference between HOV and mixed lanes (μ_2)	5.263×10^{-3}	0.194

NOTE: number of observations = 568; number of choices = 2; rho-square = 0.639; adjusted rho-square = 0.634; $L(0) = -393.71$; $L(0)$ = log likelihood function value when all parameters are 0; $L(b) = -142.07$; $L(b)$ = maximized log likelihood function value.

TABLE 3 Estimated Results of Fixed Empirical CPT-Based Logit Model

Variable	Parameters	t-Statistics
HOV alternative specific constant (μ_1)	−2.756	5.604
Speed difference between HOV and mixed lanes (μ_2)	9.925×10^{-3}	0.323

NOTE: number of observations = 568; number of choices = 2; rho-square = 0.639; adjusted rho-square = 0.634; $L(0) = -393.71$; $L(0)$ = log likelihood function value when all parameters are 0; $L(b) = -142.04$; $L(b)$ = maximized log likelihood function value.

with root-mean-squared error (RMSE) are summarized for the three models in the following table:

Model	*Observed % HOV Switch*	*Estimated % HOV Switch*	*RMSE (%)*
Estimated CPT	4.07	7.88	20.13
Expected Utility	4.07	7.06	19.97
Fixed CPT	4.07	7.15	19.97

The RMSE indicates that the estimated CPT model has similar results to those of the other two models, which lends weight to the validity of the GA approach and results for the estimates of the CPT model. This experiment is not designed to prove that one model is better than another, and such a claim is not made here.

Instead, the focus is on the estimated CPT parameters using GA, which resulted in the most statistically significant coefficients compared to the other two approaches. The estimated parameters offer new insights to how prospect framing and risk attitudes can be similar and yet vastly different for drivers on a roadway making quick decisions to get to their destinations. The similarity of the $\hat{\alpha}$ and $\hat{\beta}$ values to the empirical values suggest similar sensitivities in the value function. In addition, the estimated value $\hat{\lambda}$ is slightly higher than the empirical value (2.394 versus 2.25), which agrees with the discussions by Avineri (*12*) and de Borger and Fosgerau (*14*) about the higher loss aversion for travel time. This lends more credence to the estimated results.

However, it is interesting to see that the estimated values of γ and δ for the weight functions differ significantly from the empirical values. Both cases significantly undervalue the probability of occurrence, which means a 50% chance of occurrence may be perceived as 5% or less by a driver.

This could be due to an inability of drivers to realistically capture speed distributions at a point in time, so that there is a much greater presence of the certainty effect [or, as described by Senbil and Kitamura (*11*), a subcertainty effect in which the sum of the perceived probabilities will be much less than 1] in this experiment. For example, a well-spread distribution of many vehicles at different speeds will be perceived less strongly than many vehicles going at the same speed interval. This makes intuitive sense because the driver needs to make quick decisions on the road, and most people can evaluate the speeds of large platoons of vehicles better than they can many individual vehicle speeds.

Essentially, the distortions in the estimated parameters of the weight function suggest that the fast-paced environment of real-time traffic flow can lead to the certainty effect overwhelming the risk attitudes of drivers. A graphical illustration is shown in Figure 1 for the $\hat{\delta}$ value, which is used for calibrating the perceived probability of relative losses as demonstrated in Equation 4*b* (for the $\hat{\gamma}$ value the results are all nearly at 0 except at $p = 1$).

Figure 1*a* shows the intended mapping of the real probability on the x-axis to the perceived probability for relative losses based on calibrations from the interviews conducted by Kahneman and Tversky. For example, a decision maker facing a 50% probability of losing utility from his current position would perceive that as a slightly lower value, closer to 45%. However, Figure 1*b* shows the estimated mapping of the real probability to the perceived probability obtained for the drivers in the experiment for HOV lane selection. In this scenario, a driver facing a losing utility with 50% real probability would perceive it as less than 10%. In fact, the mapping is much more nonlinear than the empirical calibration.

This result may be interpreted as a driver who is ignoring low-probability frames. In the experiment for HOV lane selection, the driver has a relatively short period in which to make her decision to switch to the HOV lane. Because of the time constraint, she may need to perceive only a budgeted allowance of frames. As the time constraint is relaxed, the driver would tend toward the empirical values. On further speculation, this concept may be used to incorporate reaction time in prospect theory; by defining $\hat{\delta} \equiv \hat{\delta}(t)$ and $\hat{\gamma} \equiv \hat{\gamma}(t)$, a

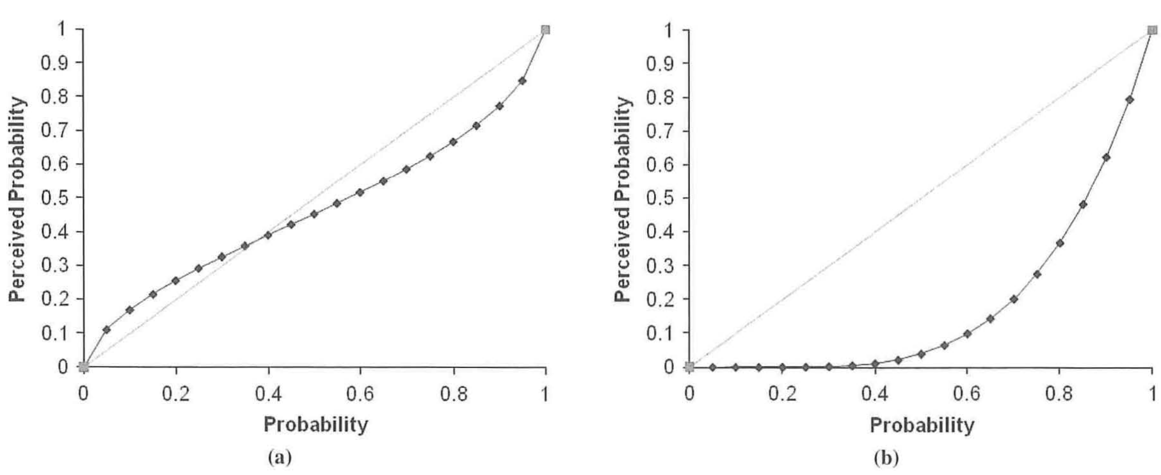


FIGURE 1 Comparison of δ: (*a*) empirical versus (*b*) GA estimated.

reaction time τ may be incorporated such that there exists a multiplier ϵ, where $\hat{\delta}(\in\tau) = 0.69$ and $\hat{\gamma}(\in\tau) = 0.61$. Stricter constraints in decision-making time would then lead to distortions in the parameters, such as those witnessed in estimated values in the HOV experiment. Future research should investigate this dependency in the cumulative prospect theory parameters.

Implications for Travel Behavioral Modeling

The experiment conducted on HOV lane selection highlights several implications that need to be considered in travel behavioral modeling. First, frame dependency and risk aversion using the static empirical calibrations may not be sufficient; the evidence suggests that the parameters may depend on different traffic conditions or, as speculated, reaction-time dependencies.

The model results from the fixed CPT and EU are very similar in likelihood and validation. This conclusion argues against the use of fixed CPT with the empirical parameters in favor of the simpler EU.

Several traveler choices have been modeled with prospect theory in the literature, including departure time, route choice, and mode choice. For practical purposes, these traveler choices have been implemented in various simulation packages that may offer some descriptive behavior. They need to be reevaluated with consideration that there may be traffic or reaction-time dependencies in the parameters leading to a dynamic descriptive behavioral model.

Implications to Traveler Choice Policies

In the real world, the impact of ignoring the dynamic dependencies of the descriptive models may be minimal for some traveler models. For example, the choice of departure time in general is not made in an environment that would suggest strong reaction-time dependencies in the parameters.

However, some traveler choice models fall under this category of a highly dynamic traffic environment with tight time constraints in decision making. From a practical point of view, any policy that requires microscopic simulation or dynamic traffic assignment to fully capture the consequences and impact would fit into this category. Examples are implementing an advanced traveler information system on a road network and coordinating disaster evacuations. For these situations, it would be crucial to incorporate realistic traveler behavior, and a dynamic descriptive behavior model needs to accompany the dynamic traffic modeling.

CONCLUSIONS

Recent literature suggests a need for a more-realistic representation of driver behavior. Clearly there are practical applications that need such considerations, such as dynamic toll pricing or disaster evacuation, where drivers' risk attitudes and dynamic choices are crucial. To integrate prospect theory as a potential descriptive method into traveler behavior models, it is important to determine the validity of previous estimated parameters obtained from empirical studies and interviews of people who are not in the same fast-paced, dynamic, travel time disutility setting as real drivers would be. This research used simplified traveler choice models in a basic scenario of HOV lane selection to test the validity of those empirical parameters. The findings suggest that there is an additional layer of complexity to perceived probabilities of drivers, supporting the hypothesis that there may be a dynamic component to prospect theory that has not yet been explored.

It has been shown that simultaneously estimating multiple parameters by using closed-form derivative-based methods can lead to estimation problems. Therefore, a heuristic GA was used to simultaneously estimate the parameters of the CPT value and weight functions as well as the coefficients of the random utility model, leading to estimates that have a higher likelihood value and more statistically significant coefficients than an equivalent EU-based logit model or a CPT-based logit model using the empirical values developed earlier. Validation with another set of data in a similar time period indicates that the GA-estimated CPT model can serve as a sufficient substitute in the experiment over the other two models for accuracy.

The estimated parameter values lead to several significant insights. The value function parameters generally conform to conclusions from previous literature. The weight function parameters, on the other hand, suggest that drivers in a fast-paced changing environment with multiple subjects for prospect evaluation may become overwhelmed by the certainty effect.

The results of the experiment suggest that there may be a reaction-time dependency in some of the prospect theory parameters. Quicker decision making is one possible reason for stronger subcertainty effects witnessed in the experiment. Further studies should be conducted to verify these new findings under different traffic speed or congestion settings, in particular with regard to the hypothesis that tighter time constraints to decision making such as faster traffic flow conditions can lead to greater distortions in the parameters. Practical applications extending from this work and Avineri's revised stochastic user equilibrium (*12*) can have tremendous benefits for evaluating route choice and traffic assignment under great uncertainty and traveler risk, such as dynamic traffic assignment under disaster evacuation.

REFERENCES

1. Simon, H. A. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics,* Vol. 69, No. 1, 1955, pp. 99–118.
2. Kahneman, D., and A. Tversky. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica,* Vol. 47, No. 2, 1979, pp. 263–291.
3. Mahmassani, H. S., and G. L. Chang. On Boundedly Rational User Equilibrium in Transportation Systems. *Transportation Science,* Vol. 21, No. 2, 1987, pp. 89–100.
4. Bonsall, P. Traveller Behavior: Decision-Making in an Unpredictable World. *ITS Journal,* Vol. 8, No. 1, 2004, pp. 45–60.
5. de Palma, A., and N. Picard. Route Choice Decision Under Travel Time Uncertainty. *Transportation Research Part A,* Vol. 39, No. 4, 2005, pp. 295–324.
6. Chen, A., Z. Ji, and W. Recker. Travel Time Reliability with Risk-Sensitive Travelers. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1783,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 27–33.
7. Ben-Akiva, M., D. McFadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou, and V. Rao. Extended Framework for Modeling Choice Behavior. *Marketing Letters,* Vol. 10, No. 3, 1999, pp. 187–203.
8. Hess, S., J. M. Rose, and D. A. Hensher. Asymmetric Preference Formation in Willingness to Pay Estimates in Discrete Choice Models. *Transportation Research Part E,* Vol. 44, No. 5, 2008, pp. 847–863.
9. Avineri, E. A Cumulative Prospect Theory Approach to Passenger Behavior Modeling: Waiting Time Paradox Revisited. *Journal of Intelligent Transportation Systems,* Vol. 8, No. 4, 2004, pp. 195–204.

10. Avineri, E., and J. N. Prashker. Sensitivity to Travel Time Variability: Travelers' Learning Perspective. *Transportation Research Part C*, Vol. 13, No. 2, 2005, pp. 157–183.
11. Senbil, M., and R. Kitamura. Reference Points in Commuter Departure Time Choice: A Prospect Theoretic Test of Alternative Decision Frames. *Journal of Intelligent Transportation Systems,* Vol. 8, No. 1, 2004, pp. 19–31.
12. Avineri, E. The Effect of Reference Point on Stochastic Network Equilibrium. *Transportation Science,* Vol. 40, No. 4, 2006, pp. 409–420.
13. Avineri, E., and P. H. L. Bovy. Identification of Parameters for a Prospect Theory Model for Travel Choice Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2082,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 141–147.
14. de Borger, B., and M. Fosgerau. The Trade-Off Between Money and Travel Time: A Test of the Theory of Reference-Dependent Preferences. *Journal of Urban Economics,* Vol. 64, No. 1, 2008, pp. 101–115.
15. Tversky, A., and D. Kahneman. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty,* Vol. 5, No. 4, 1992, pp. 297–323.
16. Ghosh, A. *Valuing Time and Reliability: Commuters' Mode Choice from a Real Time Congestion Pricing Experiment.* PhD dissertation. Department of Economics, University of California, Irvine, 2001.
17. Yan, J., K. A. Small, and E. C. Sullivan. Choice Models of Route, Occupancy, and Time of Day with Value-Priced Tolls. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1812,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 69–77.
18. Dahlgren, J. High-Occupancy/Toll Lanes: Where Should They be Implemented? *Transportation Research Part A,* Vol. 36, No. 3, 2002, pp. 239–255.
19. Choudhury, C. F. *Modeling Lane-Changing Behavior in Presence of Exclusive Lanes.* MS thesis. Massachusetts Institute of Technology, Cambridge, 2005.
20. NGSIM. FHWA, U.S. Department of Transportation. http://www.ngsim.fhwa.dot.gov/. Accessed Apr. 2008.
21. Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis.* MIT Press, New York, 1985.
22. Gen, M., and R. Cheng. *Genetic Algorithms and Engineering Optimization.* John Wiley and Sons, New York, 2000.
23. Eshelman, L., and J. Schaffer (1993). Real-coded genetic algorithms and interval-schemata. *Foundation of Genetic Algorithms,* Morgan Kaufmann Publishers, San Francisco, Cal., pp. 187–202.
24. Holland, J. H. *Adaptation in Natural and Artificial Systems.* University of Michigan Press, Ann Arbor, 1975.
25. Rennard, J. P. *Introduction to Genetic Algorithms.* http://www.rennard.org/alife/english/gavintrgb.html. Accessed July 1, 2008.
26. *GAUSS Manual.* Aptech Systems. Maple Valley, Wash., 1994.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Analyzing Route Choice Behavior with Mobile Phone Trajectories

Johannes Schlaich

Drivers have available an increasing number of options for informing themselves about the current traffic state and to subsequently make route choice decisions based on this information. Governments are investing heavily to build and operate variable message signs (VMS), drivers are buying navigation systems that show the current traffic state, and third-party providers are setting up systems that give information about the traffic state via Internet or phone. To evaluate public investments particularly, it is necessary to know how this information affects route choice behavior. Analyses were made of route choices in two study areas equipped with VMS that give recommendations for routes. The influences of VMS, broadcast traffic state, and other factors are evaluated by using maximum likelihood estimations. A unique feature of the analyses is the large database of observed route choice decisions, which are extracted from mobile phone trajectories. The total number of observations analyzed for this paper amounts to 1.4 million.

Drivers often have a choice between a number of alternative routes to their destination. However, some drivers always use the same route for a specific origin–destination pair. Others decide on their route before the trip, based on their (time-of-day) network knowledge or based on real-time information (e.g., from the Internet). Another group of drivers start their trip and are flexible to divert from their planned route on the basis of real-time information about traffic state.

Because of a higher availability of on-trip information, for example, through navigation systems or via variable message signs (VMS) along the route, a larger number of drivers could be influenced by real-time information. An optimal diversion of drivers to routes with low saturation allows road-infrastructure operators to reduce congestion, total travel times, and emissions by utilizing the existing infrastructure more efficiently. Governments see the potential in this approach; in Germany, 80 million euros were assigned to a single investment program to build new VMS systems with route recommendations on highways between 2002 and 2007 (*1*).

As emphasized by Bonsall (*2*), it is necessary to know how drivers react to this information in order to evaluate effects on the road network. However, it is not possible to observe route choice behavior with roadside detectors such as loops, because these detectors do not allow for the recognition of vehicles at different measurement points. This is why many researchers derive information from surveys by using stated preference (SP) methods with relatively low sample sizes.

Universitaet Stuttgart, Pfaffenwaldring 7, D-70569 Stuttgart, Germany. johannes.schlaich@isv.uni-stuttgart.de.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 78–85.
DOI: 10.3141/2157-10

A fairly new data source for estimation of the traffic state is floating phone data (FPD). In addition to traffic-state estimation, FPD allows for the generation of vast numbers of time–space trajectories on the road network (*3–6*). The method, which was used to generate the trajectories for the analyses in this paper, is described briefly by Friedrich et al. (*7*) and in more detail by Schlaich et al. (*8*). For this method, it is not necessary to select specific users of the mobile phone network because the data needed for the generation of FPD trajectories are available from existing equipment during the regular operation of the mobile infrastructure. The method generates trajectories based on the so-called location area updates, which are done also in the standby mode of handsets. Thus, for all users of a specific mobile network, trajectories from their origins to their destinations can be anonymously generated, not just trip segments covered during a phone call. However, only trips during 1 day can be assigned to the same user, leading to the undesirable fact that long-term analyses of specific users are not possible.

This paper reports on route-choice analyses by using the FPD trajectories of 86 days between June and October 2008 in two study areas equipped with VMS in southeast Germany:

Case I. Motorway quadrangle (Motorways A5, A6, A8, A81) and
Case II. Stuttgart (Motorways A81 and Trunk Roads B10, B27, B295).

The floating phone data were provided by T-Mobile Germany, which with some 40 million customers has a market share of roughly 36%.

Detailed analyses with sample sizes of more than 1 million observed route choice decisions will show the impact of different attributes on route choice. The paper will conclude with some considerations about the effectiveness of the examined VMS and the applied methodology using FPD trajectories as revealed-preference (RP) data.

LITERATURE REVIEW

SP experiments are the main source for the analyses of dynamic route choice behavior. Analyses with RP data are rare.

For SP experiments, it is important that the respondents are presented with a realistic context when deciding between different routes. For that reason, some SP experiments use laboratories with driving simulators and offer financial rewards to those respondents who arrive quickly at their destination, to ensure serious responses (*9*). The financial reward should imitate the bonus of earlier arrival when making a good route choice in real life.

The results of different studies often vary significantly, although they may use the same or very similar approaches. For example,

Wardman et al. (*10*) and Kim and Chon (*11*) found different results for the impact of VMS. As done by Kim and Chon, many discrepancies can be explained by hypotheses about slightly different research designs. However, some of the differences may result from different layouts of the SP experiment. For example, different layouts of questionnaires may influence respondents. A nicely presented VMS may be more trustworthy for respondents than a photograph of an old VMS. Also, different ways of recruiting respondents may influence the results significantly. Undergraduate students chasing a credit (*12*) will probably provide different answers than will respondents recruited at highway petrol stations (*13*).

However, RP analyses often face significant problems. First, the recruitment of people to fill in questionnaires or carry Global Positioning System devices is difficult (*14*), often resulting in skewed results. Second, when questionnaires are used to determine trips, there often is great inaccuracy regarding the exact departure time and chosen route. Third, in analysis of choice behavior on the basis of RP data, external factors cannot be influenced in real life, and this problem can result in a large number of observations without active VMS or broadcast traffic news.

Detailed results of previous SP and RP studies are available elsewhere (*9–17*).

MATHEMATICAL BACKGROUND

Readers who are familiar with the concepts of the utility function, logit model, and maximum likelihood estimation may skip this section.

Utility Function

Route choice models usually need a utility function, which combines different relevant attributes into a final utility for each alternative. A utility function often consists of an alternative-specific constant parameter α and different terms for each attribute A_n, which influences route choice, and a corresponding β_n parameter:

$$V_j = \alpha_j + \sum_{n=1}^{N} \beta_n \cdot A_{n,j}$$

where

V_j = deterministic utility of alternative j,
N = number of attributes,
$A_{n,j}$ = nth attribute of alternative j, and
α_j, β_n = parameters.

Ben-Akiva et al. list, among others, time, distance, road quality, and hierarchical travel pattern as potential attributes A_n (*18*).

Logit Model

The logit model is a distribution model that assigns traffic volumes of specific origin–destination pairs to alternative routes for the same origin–destination pairs on the basis of the differences in utility values:

$$P_j = \frac{e^{V_j}}{\sum_{n=1}^{N} e^{V_j}}$$

where

P_j = probability of alternative j,
V_j = deterministic utility of alternative j, and
N = number of alternatives.

Maximum Likelihood Estimation

The α and β parameters of the utility function are a priori unknown. However, they can be estimated by using SP or RP data. Use of maximum likelihood estimation maximizes the following function $L(\alpha_i, \beta_i)$ by optimizing the α and β parameters:

$$L(\alpha_j, \beta_n) = \sum_{k=1}^{K} \sum_{j=1}^{J} g_{k,j} \cdot \ln P\left(V\left(\alpha_j, \beta_n\right)\right)_{k,j}$$

where

$L(\alpha_j, \beta_n)$ = log likelihood function with parameters α_i and β_i,
$k = 1$ to K = number of observed decisions,
$j = 1$ to J = number of alternatives for a decision k,
$P(V)_{k,j}$ = probability of alternative j for a decision k,
$V(\alpha_j, \beta_n)_{k,j}$ = utility of the alternative j for the observation k, and
$g_{k,j}$ = choice variable: 1 if alternative j for observation k is chosen and 0 if alternative j for observation k is not chosen.

The quality of the estimation can be evaluated with the likelihood ratio test, which compares the values of the log likelihood function of different estimations:

$$\text{LR} = 2 \cdot (L_2 - L_1)$$

where LR is the likelihood ratio and L_1, L_2 are values of the log likelihood function of different estimations, where L_2 is a restricted version of L_1.

The likelihood ratio follows approximately a chi-squared distribution (*19*). Thus, the significance of the results can be checked by using the number of additional parameters in the two estimations as the degree of freedom. In addition, the significance of the estimated parameters should be controlled by using t-statistics.

The free software package Biogeme 1.8 was used for the estimations in this paper. Information about the software and further useful references regarding the maximum likelihood estimation are available elsewhere (*20, 21*).

CASE I. ROUTE CHOICE BEHAVIOR IN MOTORWAY QUADRANGLE

Case I Study Area and Sample

In the motorway quadrangle study area (see Figure 1), two alternatives are available along both diagonals:

- Stuttgart ↔ Walldorf diagonal
  - Stuttgart–Karlsruhe–Walldorf (S-KA-W), 91 km
  - Stuttgart–Heilbronn–Walldorf (S-HN-W), 103 km
- Karlsruhe ↔ Heilbronn diagonal
  - Karlsruhe–Walldorf–Heilbronn (KA-W-HN), 92 km
  - Karlsruhe–Stuttgart–Heilbronn (KA-S-HN): 102 km

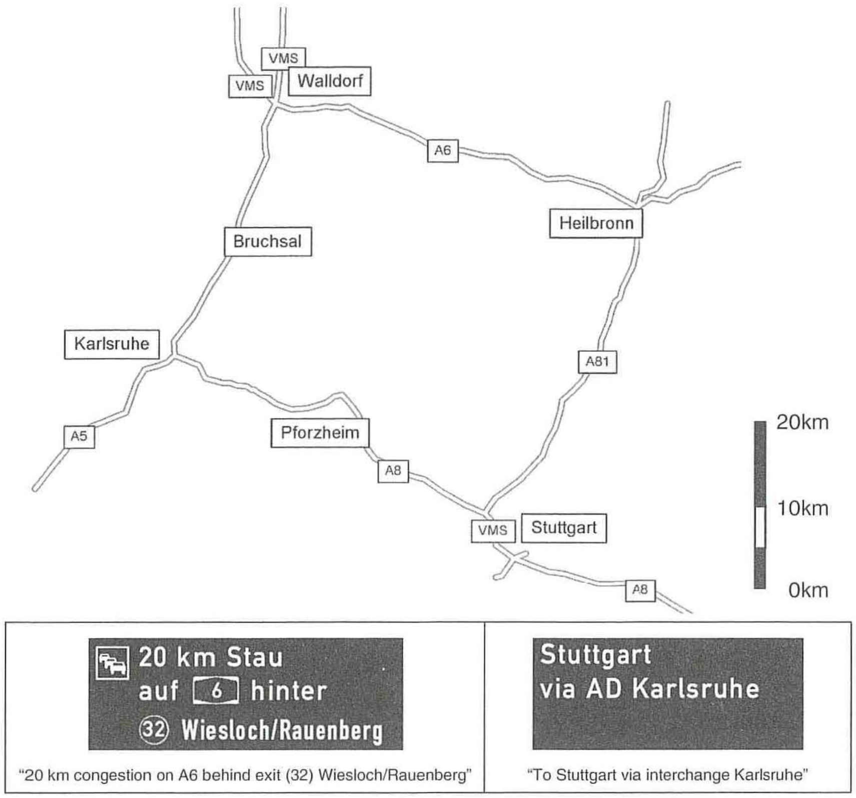


FIGURE 1 Motorway quadrangle study area and VMS.

For the Stuttgart ↔ Walldorf journey, static signs recommend the longer route via Heilbronn, but VMS recommend driving via Karlsruhe in the case of congestion on the Heilbronn route. The VMS show the recommendation along with information about the length and location of the congestion. Along the other diagonal, static signs recommend the route via Walldorf. There are no VMS for this diagonal.

Table 1 gives a summary of the sample for route choice analyses in the motorway quadrangle. A total of 1,059,712 trajectories are available for analysis. Roughly 70% of the trips along the Stuttgart ↔ Walldorf diagonal do not use the route recommended by the static signs. (This is reasonable, since the average travel time and distance of the route via Heilbronn are greater.)

As shown in Table 1, the FPD trajectories can be roughly divided into different vehicle classes on working days. This division is done on the basis of their individual travel times in relation to the floating average of the travel times of all vehicles (*5*).

Estimation of Basic Model

The basic model consists of constant α parameters for the most popular route in each direction as well as β parameters for the VMS recommendation (only Stuttgart ↔ Walldorf) and the traffic state broadcast via radio news and the traffic management center (TMC). In Germany, traffic state news is divided into congestion (*Stau*) and

TABLE 1 Summary of Through Traffic

Direction	Route	Through-Traffic Vehicles	Share (per direction) (%)	Share of Cars (only Mon.–Fri.) (%)	Share of Trucks (only Mon.–Fri.) (%)
S-W	S-KA-W	235,628	71.2	50.9	49.1
	S-HN-W	95,262	28.8		
W-S	W-KA-S	271,188	70.5	50.0	50.0
	W-HN-S	113,373	29.5		
Total S ↔ W		715,451			
KA-HN	KA-W-HN	160,701	95.2	48.1	51.9
	KA-S-HN	8,092	4.8		
HN-KA	HN-W-KA	169,424	96.6	50.9	49.1
	HN-S-KA	6,044	3.4		
Total KA ↔ HN		344,261			

NOTE: S = Stuttgart, KA = Karlsruhe, HN = Heilbronn, and W = Walldorf.

halting (*Stockend*) with congestion being more severe. This results in the following utility functions:

- General form of the utility function:

$$V_{j,d} = \alpha_{j,d} + \beta_{\text{VMS}} \cdot \text{VMS}_{\text{active},d} + \beta_{l\_\text{congestion}} \cdot l_{\text{congestion},j,d} + \beta_{l\_\text{halting}} \cdot l_{\text{halting},j,d}$$

- Diagonal Stuttgart ↔ Walldorf:

$$V_{\text{KA},d} = \alpha_{\text{KA},d} + \beta_{\text{VMS}} \cdot \text{VMS}_{\text{active},d} + \beta_{l\_\text{congestion}} \cdot l_{\text{congestion},\text{KA},d} + \beta_{l\_\text{halting}} \cdot l_{\text{halting},\text{KA},d}$$

$$V_{\text{HN},d} = 0 + \beta_{l\_\text{congestion}} \cdot l_{\text{congestion},\text{HN},d} + \beta_{l\_\text{halting}} \cdot l_{\text{halting},\text{HN},d}$$

- Diagonal Karlsruhe ↔ Heilbronn:

$$V_{W,d} = \alpha_{W,d} + \beta_{l\_\text{congestion}} \cdot l_{\text{congestion},W,d} + \beta_{l\_\text{halting}} \cdot l_{\text{halting},W,d}$$

$$V_{S,d} = 0 + \beta_{l\_\text{congestion}} \cdot l_{\text{congestion},S,d} + \beta_{l\_\text{halting}} \cdot l_{\text{halting},S,d}$$

where

$V_{j,d}$ = deterministic utility of alternative j in direction d;
$\alpha_{j,d}$ = constant parameter for alternative j in direction d;
β_n = parameters for attributes n (e.g., one congestion for length of congestion);
$\text{VMS}_{\text{active},d}$ = 1 if VMS gives a route recommendation in the d direction, 0 otherwise (at the time a car is passing the VMS station, this does not apply to diagonal Karlsruhe ↔ Heilbronn);
$l_{\text{congestion},j,d}$ = length of congestion for alternative j in direction d (total length of active news at the time of route choice decision);
$l_{\text{halting},j,d}$ = length of halting traffic for alternative j in direction d (total length of active news at the time of route choice decision);
KA = Karlsruhe;
HN = Heilbronn;
W = Walldorf;
S = Stuttgart; and
d = direction (e.g., S-W or W-S).

Table 2 shows highly significant results for the likelihood ratio and all estimated parameters. Furthermore, the results remain very similar when the model is estimated with only half the data, whatever half is chosen, or when estimating the β parameters per origin–destination pair.

The constants α are positive as they are assigned for the most popular alternative for each route choice. The values for the Karlsruhe ↔ Heilbronn diagonal are higher than for the other diagonal, as the distribution is with 95% of the observed trips using the main route clearer than for the other diagonal.

The β parameters for the broadcast traffic state have negative values as expected, because they constitute a negative utility for the driver. β_{l_halting} is roughly 16% lower than $\beta_{l\_\text{congestion}}$, which shows that the drivers and navigation systems using the broadcast traffic state via TMC can differentiate between the two different traffic states. For comparison, analyses by Schlaich and Friedrich for this motorway quadrangle showed that the real impact on travel time of the halting message is 37% lower than the impact of the congestion message (*16, 17*).

The value for β_{VMS} is positive, which means that drivers follow the VMS recommendation to drive via Karlsruhe. It can be concluded from a comparison of the values of β_{VMS} and $\beta_{l\_\text{congestion}}$ that a broadcast message of 7-km congestion has—other things being equal—the same impact as the recommendation of the VMS.

The route choice behavior can be plotted on a diagram such as in Figure 2 by using the estimated parameters. Such a diagram allows estimation of the impact of broadcast traffic news and VMS route recommendations. For example, if the broadcast congestion via Heilbronn is 5 km longer than congestion via Karlsruhe, a route recommendation from the VMS would increase the share of through traffic via Karlsruhe by 6.6%, from 77.7% to 84.3%. This means an acceptance of the VMS recommendation of 29.3%:

$$\text{share}_{\text{acceptance,VMS}} = \frac{\text{share}_{\text{through-traffic,diverted}}}{\text{share}_{\text{through-traffic,divertable}}} = \frac{6.6\%}{(100\% - 77.7\%)} = 29.3\%$$

According to this scenario, a mere 0.79% of the total traffic volume at the VMS is diverted, because the share of the through traffic is 12% (*16, 17*):

$$q_{\text{divereted,VMS}} = q_{\text{total,VMS}} \cdot \text{share}_{\text{through-traffic}} \cdot \text{share}_{\text{through-traffic,diverted}} = q_{\text{total,VMS}} \cdot 12\% \cdot 6.6\% = q_{\text{total,VMS}} \cdot 0.79\%$$

For a total traffic volume at the VMS of, for example, 6,000 vehicles per hour, this leads to fewer than 50 diverted vehicles per hour.

TABLE 2 Results of Basic Model in Motorway Quadrangle

Parameter		Results			
No.	Name	Value	SE	t-Statistic	Description
1	$\alpha_{\text{KA,S-W}}$	0.936	0.021	45.188	Constants for the preferred route of each route choice decision
2	$\alpha_{\text{KA,W-S}}$	0.924	0.021	43.886	
3	$\alpha_{\text{W,KA-HN}}$	2.968	0.048	61.702	
4	$\alpha_{\text{W,HN-KA}}$	3.328	0.057	58.734	
5	$\beta_{l\_\text{congestion}}$	−0.062	0.003	−19.929	Parameters for the broadcast traffic state
6	β_{l_halting}	−0.052	0.005	−11.624	
7	β_{VMS}	0.428	0.070	6.079	Parameter for the VMS

NOTE: Total observations, 1,059,712; null log likelihood, −734,431.6; log likelihood, −486,300.8; likelihood ratio (LR), 496,261.7; adjusted ρ^2, 0.3375.

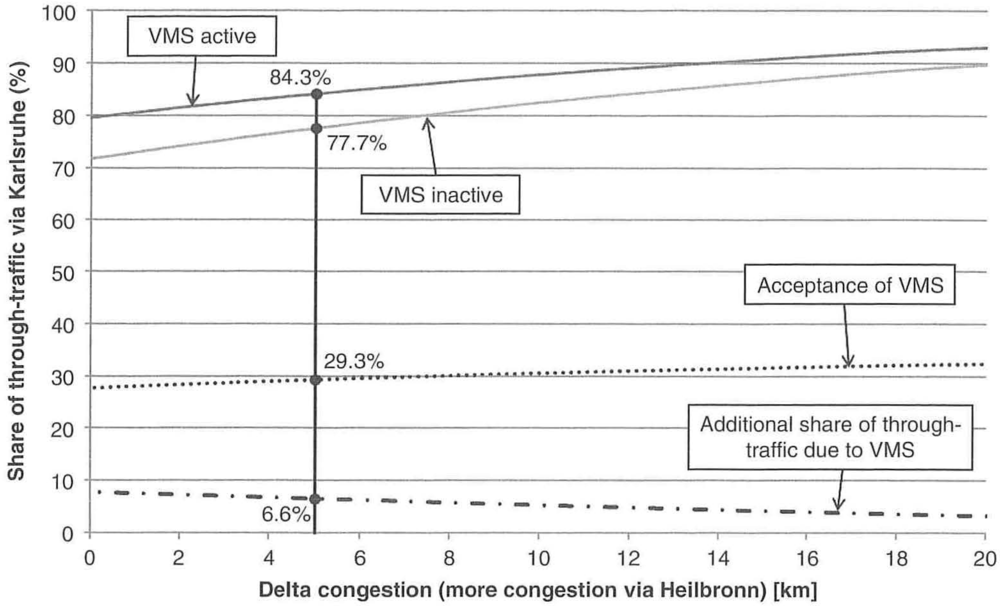


FIGURE 2 Route choice behavior for Stuttgart ↔ Walldorf diagonal when there is more congestion via Heilbronn than via Karlsruhe.

Accounting for Location of Broadcast Traffic Congestion

In a further estimation, the two β parameters for the broadcast traffic state will be divided into four parameters by considering the location of the broadcast traffic state. If the location of congestion or halting traffic is on the first edge of the route in the motorway quadrangle, it will be assigned to $\beta_{l\_congestion/halting,1st\,half}$, and if it is on the second edge, it will be assigned to $\beta_{l\_congestion/halting,2nd\,half}$.

Table 3 shows the results of this estimation, which are highly significant with a likelihood ratio of 887.5 compared to the basic model. The β parameters for the first half are 70% to 90% higher than for the second half of the route. This means that the drivers (or their navigation systems) more highly value this broadcast traffic news, which they will reach earlier, than that which they will reach later. This is because drivers expect the farther congestion with a higher probability to have disappeared before their arrival. The estimation shows that the drivers and navigation systems are capable of processing this differentiated data.

Further Extensions of Basic Model

The large sample size enables several further detailed estimations:

- A differentiation between cars and trucks shows similar route choice behavior for both vehicle classes. Trucks, which pay a distance-dependent toll, tend to use the shorter route a little more. Cars respond more to the broadcast traffic state. However, the total fit of the model, represented by the adjusted-ρ^2, is not much better than the basic model.
- Replacing the alternative-specific constants α by the attribute "typical time-of-the-day and day-of-the-week travel time" from long-time measurements results in a worse model fit. This indicates

TABLE 3 Accounting for Location of Broadcast Traffic News

Parameter		Results			
No.	Name	Value	SE	*t*-Statistic	Description
1	$\alpha_{KA,S\text{-}W}$	0.951	0.021	45.390	Constants for the preferred route of each route choice decision
2	$\alpha_{KA,W\text{-}S}$	0.873	0.025	35.390	
3	$\alpha_{W,KA\text{-}HN}$	2.917	0.050	58.014	
4	$\alpha_{W,HN\text{-}KA}$	3.334	0.057	58.959	
5	$\beta_{l\_congestion,1st\,half}$	−0.080	0.004	−19.813	Parameters for the broadcast traffic state
6	$\beta_{l\_congestion,\,2nd\,half}$	−0.042	0.004	−9.344	
7	$\beta_{l\_halting,\,1st\,half}$	−0.073	0.007	−9.738	
8	$\beta_{l\_halting,\,2nd\,half}$	−0.042	0.006	−7.576	
9	β_{VMS}	0.395	0.070	5.601	Parameter for the VMS

NOTE: Total observations, 1,059,712; null log likelihood, −734,431.6; log likelihood, −485,862.0; likelihood ratio (LR), 497,139.2; LR compared to basic model, 877.5; adjusted ρ^2, 0.3384.

that drivers have reasons other than typical travel time when choosing an alternative. The parameter for typical travel time in minutes is (for the sake of completeness) roughly –0.22.

- Using the preceding typical travel time in addition to the constants does not considerably improve the model fit. It can be concluded that drivers have little knowledge about the typical travel times that they can expect. They do, however, know about the general advantage of the faster routes, which is reflected by the positive values of the constants for the faster routes.
- Drivers do not react to additional information on the VMS. The inclusion of length of displayed congestion or displayed location of congestion does not improve the model fit noticeably, although the correlation of the variables is relatively low. Thus, drivers who follow the recommendation of the VMS do this mostly regardless of the displayed additional information. This does not contradict the results of other recent publications (*22*), which say that additional information together with route information improve the acceptance of VMS. It means only that the information itself is not important.

CASE II. ROUTE CHOICE BEHAVIOR IN STUTTGART

Case II Study Area and Sample

In the Stuttgart study area (see Figure 3), there are three alternatives available from the Stuttgart center northbound in the direction of Heilbronn, and the same alternatives are available in the opposite direction:

- B10: standard route, recommended by static signs;
- B295: route recommended by VMS, about 6 km and 10 min longer than B10; and
- B27: same distance as B10, but about 6 min longer.

TABLE 4 Summary of Through Traffic in Stuttgart

Direction	Route	Number of Observations	Share (per direction) (%)
Northbound	B10	160,176	93.8
	B295	4,518	2.6
	B27	5,984	3.5
Total		170,678	
Southbound	B10	164,826	96.6
	B295	8,004	4.7
	B27	5,984	3.5
Total		178,814	

The VMS in the southbound direction into Stuttgart is situated after the exit to the B27 and thus cannot be seen by drivers taking this alternative. In the northbound direction, the driver of all alternatives can see the VMS.

Table 4 shows the sample size for the Stuttgart study area. The main route, B10, has a share of more than 90%.

Route Choice Analyses

The maximum likelihood estimation of the basic model, taking into account VMS and traffic news, does not result in significant or reasonable parameters. Table 5 shows that all β parameters are not significant according to the *t*-statistics. In addition, the parameters for the traffic news show an unexpected algebraic sign with $\beta_{l\_congestion}$ being positive. This leads to the hypotheses that traffic news has little or no effect on route choice for the Stuttgart study area, which is used mainly by regular drivers. This may be because the B27 and B295 routes, especially, are rarely covered

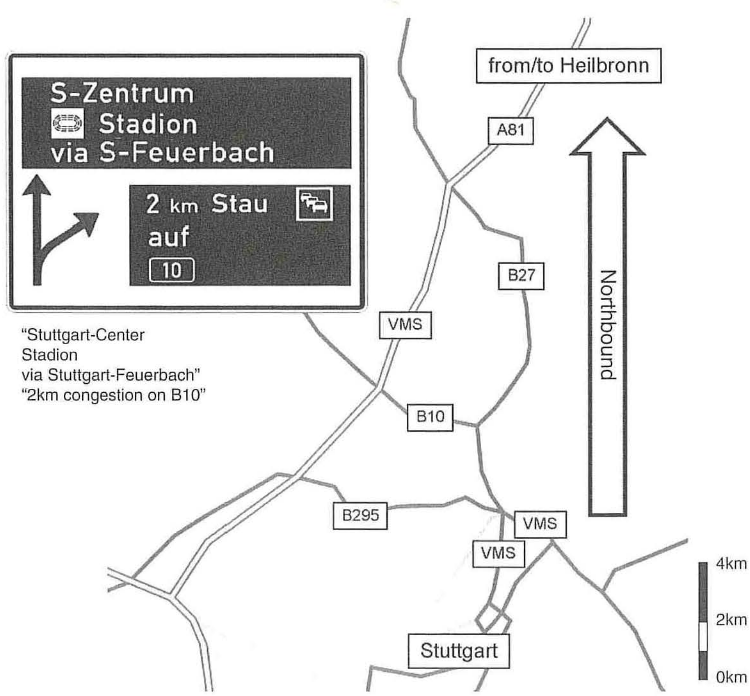


FIGURE 3 Stuttgart study area and VMS.

TABLE 5 Results of Basic Model in Stuttgart

Parameter		Results			
No.	Name	Value	SE	t-Statistic	Description
1	α_{B295}	−3.270	0.253	−12.942	Constants for the nonstandard routes of each route choice decision
2	α_{B27}	−3.304	0.224	−14.785	
3	$\beta_{l\_congestion}$	0.154	0.155	0.994	Parameters for the broadcast traffic state
4	$\beta_{l\_halting}$	−0.071	0.120	−0.587	
5	β_{VMS}	0.654	0.549	1.192	Parameter for the VMS

NOTE: Total observations, 349,492; null log likelihood, −383,956.2; log likelihood, −105,586.9; likelihood ratio, 556,738.5; adjusted ρ^2, 0.7250.

by the broadcast traffic state because of their low ranking in the road network.

The parameter β_{VMS} has the anticipated positive value. This means that drivers follow the recommendation of the VMS. However, as stated earlier, the t-statistic does not show significant result, with the standard error being nearly as high as the value of β_{VMS}. This is mainly because of the low number of observations during an active VMS. Especially in the southbound direction, the VMS was active only on six of 86 analyzed days, with an average active period of 45 min per day. Figure 4 shows the route choice behavior as a function of the VMS in the two directions. The values N below the columns display the number of observations in each column.

In the northbound direction, the VMS route recommendation to use the B295 causes only a minor increase in through traffic on that route. There is also a positive effect on B27 usage, either because of information from the VMS about congestion on the B10 or because of actual observable congestion on the first part of the B10 route, which is common with the B27 route. In the southbound direction, the effect of the VMS on through traffic is higher. The share of the B27 remains stable because the drivers using the B27 do not pass the VMS. The acceptances of the VMS are 2.7% (northbound) and 17.3% (southbound), respectively, after traffic on the B27 is disregarded.

CONCLUSIONS

This paper presented an application of FPD trajectories. Attributes influencing route choice behavior were examined for two study areas equipped with VMS.

The findings for route choice behavior are statistically significant for the motorway quadrangle:

- Drivers react to traffic news broadcast via radio news and TMC. For each kilometer of congestion on one of the two 100-km-long alternatives, on average 1% of the through traffic will change to the other route.
- Drivers analyze the location of the broadcast traffic news. The nearer the congestion, the greater the impact of the broadcast traffic news.

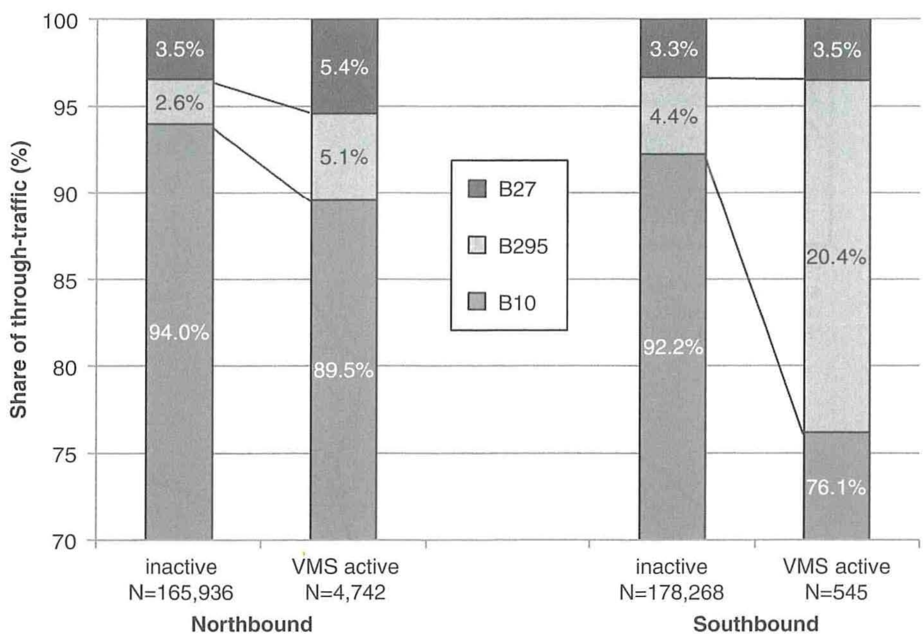


FIGURE 4 Route choice behavior for Stuttgart study area.

- Drivers react to VMS route recommendations. Thirty percent of through-traffic drivers who normally use the standard route change to the recommended route. This leads to a maximum of fewer than 100 diverted vehicles per hour.
- Drivers do not analyze additional information on the VMS about the reason for the route recommendation (here: location and length of congestion).
- Trucks and cars display similar route choice behavior. However, trucks prefer the shorter routes, whereas cars react more to dynamic information.

The analyses for the Stuttgart study area did not result in statistically significant findings; however, the acceptance of VMS route recommendations was only between 3% and 17%.

The methodology using FPD trajectories as RP data for route choice analyses showed strengths and weaknesses. A large database from even just one infrastructure operator is powerful because the incidence of mobile phones in vehicles is very high. A weakness of the FPD trajectories approach is that, except for the separation between cars and trucks, no information about the driver or equipment such as navigation systems and trip purpose is available. Another weakness regarding RP data in general showed up in analysis of the Stuttgart study area, as the analyzed VMS was seldom active, leading to statistically insignificant results.

Regardless of the few methodological drawbacks, the results of the route choice analyses using FPD trajectories are promising and useful for the design and operation of VMS. Only a detailed knowledge of the impact of VMS enables optimal usage of the capacity of the existing infrastructure. Further research with nationwide FPD trajectories would allow comprehensive RP data to be collected for several route choice decisions with different layouts, detour factors, and route lengths. Furthermore, nationwide FPD trajectories would facilitate investigations, whether constants can be replaced or at least diminished by adding appropriate attributes such as route lengths, travel time, or reliability.

ACKNOWLEDGMENT

This research was part of the Do-iT project, which was funded by the German Federal Ministry of Economy and Technology.

REFERENCES

1. *Programm zur Verkehrsbeeinflussung auf Bundesautobahnen, 2002 bis 2007.* Bundesministerium für Verkehr, Bau- und Wohnungswesen, Berlin, 2002.
2. Bonsall, P. The Influence of Route Guidance Advice on Route Choice in Urban Networks. *Transportation,* Vol. 19, No. 1, 1992, pp. 1–23.
3. Qiu, Z.-J., J. Jin, P. Cheng, and B. Ran. State of the Art and Practice: Cellular Probe Technology Applied in Advanced Traveler Information Systems. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
4. Caceres, N., J. P. Wideberg, and F. G. Benitez. Review of Traffic Data Estimations Extracted from Cellular Networks. *Intelligent Transport Systems,* Vol. 2, No. 1, 2008, pp. 179–192.
5. Schlaich, J. *Nutzung von Mobilfunkdaten für die Analyse der Routenwahl.* PhD dissertation. Department of Traffic Engineering and Transportation Planning, University of Stuttgart, Germany, 2009.
6. Borzacchiello, M. T., J. Steenbruggen, P. Nijkamp, and H. J. Scholten. Use of Data from Mobile Phone Networks for Transportation Applications. Presented at 89th Annual Meeting of the Transportation Research Board, Washington, D.C., 2010.
7. Friedrich, M., P. Jehlicka, T. Otterstätter, and J. Schlaich. Monitoring Travel Behaviour and Service Quality in Networks with Floating Phone Data. Presented at International Conference on Transportation Decision Making: Issues, Tools, Models and Case Studies, Venice, Italy, 2008.
8. Schlaich, J., T. Otterstätter, and M. Friedrich. Generating Trajectories from Mobile Phone Data. Presented at 89th Annual Meeting of the Transportation Research Board, Washington, D.C., 2010.
9. Selten, R., T. Chmura, T. Pitz, J. Wahle, and M. Schreckenberg. Commuters Route Choice Behaviour. *Games and Economic Behavior,* Vol. 58, No. 2, 2007, pp. 394–406.
10. Wardman, M., P. W. Bonsall, and J. D. Shires. Driver Response to Variable Message Signs: A Stated Preference Investigation. *Transportation Research Part C,* Vol. 5, No. 6, 1997, pp. 389–405.
11. Kim, H. R., and K. S. Chon. Modeling En-Route Diversion Behaviour Under On-Site Traffic Information. *Journal of the Eastern Asia Society for Transportation Studies,* Vol. 6, 2005, pp. 1833–1843.
12. Han, Q., H. J. P. Timmermans, B. Dellaert, and F. van Raaij. Route Choice Under Uncertainty: Effects of Recommendation. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.
13. Emmerink, R. H. M., P. Nijkamp, P. Rietveld, and J. N. van Omeren. Variable Message Signs and Radio Traffic Information: An Integrated Empirical Analysis of Drivers' Route Choice Behaviour. *Transportation Research Part A,* Vol. 30, No. 2, 1996, pp. 135–153.
14. Wermuth, M., C. Sommer, and S. Wulff. *Erhebung der individuellen Routenwahl zur Weiterentwicklung von Umlegungsmodellen.* FE-Nr. 01.158/2002/CGB. Bundesministeriums für Verkehr, Bau- und Wohnungswesen, Berlin, 2004.
15. Bierlaire, M., M. Themans, and K. W. Axhausen. Analysis of Driver's Response to Real-Time Information in Switzerland. *European Transport,* Vol. 34, 2006, pp. 21–41.
16. Schlaich, J., and M. Friedrich. Staumeldungen und Routenwahl in Autobahnnetzen. *Straßenverkehrstechnik,* No. 10, 2008, pp. 621–627.
17. Schlaich, J., and M. Friedrich. Staumeldungen und Routenwahl in Autobahnnetzen. *Straßenverkehrstechnik,* No. 11, 2008, pp. 706–712.
18. Ben-Akiva, M., M. Bergman, A. Daly, and R. Ramaswamy. Modelling Inter Urban Route Choice Behavior. *Ninth International Symposium on Transportation and Traffic Theory,* VNU Science Press, Utrecht, Netherlands, 1984.
19. Huelsenbeck, J., and K. Crandall. Phylogeny Estimation and Hypothesis Testing Using Maximum Likelihood. *Annual Review of Ecology and Systematics,* Vol. 28, 1997, pp. 437–466.
20. Bierlaire, M. Biogeme: A Free Package for the Estimation of Discrete Choice Models. *Proc., 3rd Swiss Transportation Research Conference,* Ascona, Switzerland, 2003.
21. Bierlaire, M. *An Introduction to BIOGEME Version 1.8.* http://biogeme.epfl.ch. Accessed June 8, 2009.
22. Trapp, R., and M. Feldges. Ermittlung der Wirksamkeit von Verkehrsbeeinflussungsanlagen. *Straßenverkehrstechnik,* Vol. 10, 2009, pp. 375–379.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Models of Mode Choice and Mobility Tool Ownership Beyond 2008 Fuel Prices

Claude Weis, Kay W. Axhausen, Robert Schlich, and René Zbinden

A recent project addressed how travelers would react to fuel prices rising above the high levels that were reached in mid 2008. Study participants were recruited during phone interviews, in the course of which trips made on a specified day were recorded. On the basis of one of those trips and the respondents' possession of mobility tools, stated preference (SP) experiments were constructed. The first part consisted of a mode choice situation under modified price (and travel time) settings (tactical decisions). The second part focused on long-term (strategic) choices between the current and an alternative fleet, including a redistribution of yearly mileage. From the SP data, multinomial logit models for mode and fleet choice were estimated. The mode choice models were estimated by using income- and distance-dependent nonlinear utility functions and separately for the various trip purposes (as was the practice in earlier Swiss studies on similar topics) and controlled for all relevant trip characteristics. The models for mobility tool ownership, which were formulated by using a new approach, aimed to yield trade-offs between the various attributes of the offered fleets and to forecast the distribution of annual transit passes under modified settings. The findings suggest that inertia is present in both mode choice and mobility tool ownership. Elasticities do not change much from previous studies, where more-conservative price increases were assumed. Transit pass ownership is expected to grow only when increasing fuel prices coincide with stable public transport fares.

This paper describes a study contracted by the Swiss Federal Railway Company's passenger traffic section that was recently conducted at the Institute for Transport Planning and Systems (IVT), ETH Zurich (*1*).

The first part of the study complements the results established in previous Swiss mode choice studies. Vrtic and Fröhlich presented demand elasticities derived from a comprehensive mode choice model (*2*). Axhausen et al. (*3, 4*) and Hess et al. (*5*) computed values of travel time savings for Switzerland. Their results were applied in the official Swiss cost–benefit guidelines for values of travel time savings (*6*). Literature on mobility tool—particularly transit pass—ownership is quite sparse. Most of the literature appears to focus on modeling car ownership and usage without considering other dimensions (*7–10*). This is surprising: considering transit passes as substitute goods for cars appears intuitive, at least in the context of regions with high-quality public transport service and widely available transit passes. Thus, modeling car and transit pass ownership jointly in such settings appears to be an obvious approach. Scott and Axhausen presented evidence of interactions between car and transit pass choices in a household context and show that commitment to one mode heavily influences the use of the other, confirming the substitution effect (*11*). However, their study lacks variation in the price levels, thus making it impossible to estimate their influence on mobility tool ownership. Vrtic et al. investigated the effects of potential mobility pricing schemes on travelers' tactical (mode choice) and strategic (long-term) decisions (*12, 13*). Their analyses were based on stated preference surveys with quite conservative price variations and yielded very low elasticities for transit pass ownership. Axhausen et al. used a structural equations model to test hypotheses on paths linking car ownership, transit pass ownership, and modal usage (*14*). However, they did not model the choice determinants for owning the various mobility tools.

The present study was sparked by increasing oil prices at the time it was initiated (mid 2008). Short-term mode choice and longer-term mobility tool ownership decisions were of equal interest. At the time, the price for a liter of regular nonleaded fuel was around CHF 2 (Swiss francs; as of February 2010, CHF 1 = \$0.92), a circumstance that provided an opportunity for stated preference (SP) experiments implementing a much greater bandwidth in pricing schemes than in previous studies without losing realism in the experiments. Because customer reactions are assumed to vary with costs, a mere extrapolation of previous results to the new price levels could lead to biased forecasts. Specifically, demand elasticities are expected to increase nonlinearly with rising prices. Although the current worldwide economic crisis has counteracted increases in fuel prices, fuel at CHF 4 per liter and above remain quite imaginable in the mid- to long-term future, especially in the context of fossil energy shortage scenarios. [Several works offer qualitative analyses on the effects of supply disruptions in the United Kingdom in 2000 (*15–17*).]

In the short term, the interest focused on modeling individuals' mode choice under modified pricing schemes while accounting for other relevant decision variables. Here, the figures of interest were values willingness-to-pay (WTP) indicators and elasticities. These were estimated by means of discrete choice models based on SP mode choice experiments, where the situations were based on trips reported by the recruited respondents during previous phone interviews.

In the longer term, the main interest was in how customers would adapt their sets of mobility tools, and especially transit passes, to the aforementioned changes, and what the effects in the usage of these mobility tools would be (that is, if and how a redistribution of mileage from car to public transport would take place). Therefore, a second SP questionnaire was designed, focusing on the determinants of these choices. Because an iterative and interactive approach to the

C. Weis, HIL F 33.1, ETH Hönggerberg, and K. W. Axhausen, HIL F 32.3, ETH Hönggerberg, Institute for Transport Planning and Systems (IVT), ETH Zurich, Wolfgang-Pauli-Str. 15, CH-8093 Zurich, Switzerland. R. Schlich, SBB Local Traffic, and R. Zbinden, SBB Passenger Traffic, Swiss Federal Railways, Wylerstrasse 125, CH-3000 Berne, Switzerland. Corresponding author: C. Weis, weis@ivt.baug.ethz.ch.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 86–94.
DOI: 10.3141/2157-11

respondents' most-desirable setting was not feasible (pen-and-paper questionnaires were used as the survey instrument), an approach was used in which the interviewees were given choices between their current mobility tool setting (at new price levels) and alternative fleets, consisting of a new car, a redistribution of annual mileage, and the corresponding cheapest annual transit pass. The fleet choice models were formulated as mixed multinomial logit models incorporating the two offered fleets' attributes as linear terms in the utility functions. This approach, which is fairly straightforward and, to the authors' knowledge, new in its application, was chosen over a discrete-continuous formulation because of data considerations and to understand the trade-offs between transit pass prices and their ownership and use. Under the assumptions made in the study, the results should allow revenue managers to choose pricing strategies for transit passes as a function of fuel prices.

DATA COLLECTION

Respondent Recruitment

The data used for the study result from an SP survey. Respondents were recruited through an ongoing continuous computer-assisted telephone survey, in which approximately 400 persons per week were interviewed. The survey, called the kontinuierliche Erhebung Personenverkehr (KEP), was contracted by the Swiss Federal Railway Company and conducted by Link, and it comprised questions on respondents' sociodemographic characteristics and travel behavior (*18*). Trips of more than 3 km undertaken during a week are recorded along with relevant characteristics (origin, destination, travel and waiting times, distances). Following the interview, respondents who had a driving license were recruited for the SP survey. Limiting the recruitment process to licensed drivers ensured that the car alternative presented in the mode choice experiments was available to all respondents. For those fulfilling the requirements and who were willing to participate in the study, one reported trip was selected according to the following criteria:

- If at least one car trip was reported, the longest such trip was selected.
- Otherwise, the longest reported trip by railway was selected.

The personalized mode choice experiments for each respondent were constructed from the selected trip. This procedure made sure that every participant was presented with choice situations tailored to their actual behavior instead of hypothetical scenarios.

From September 8 to November 9, 2008 (calendar weeks 38 through 45), 1,200 respondents were recruited for the study. For various reasons, 200 of the recruited persons were excluded from the final survey; thus there was a total sample size of 993 participants, to which questionnaires were sent by post.

Design of SP Experiments

Mode Choice

In accordance with current practice (*19–23*), the SP experiments were designed on the basis of data reported by the respondents in the phone interviews. This procedure has been successfully applied in former Swiss studies (*12, 13, 24, 25*). The attribute levels used in the mode choice SP experiment were derived from the chosen trip. In construction of the experiments, the existing attributes of the trip and its mode alternative were increased or decreased by predetermined factors, which are shown in Table 1.

The attributes for the car alternative were derived from the Swiss network model (*26*). For the rail alternative, they were obtained from the Swiss Federal Railway Internet timetable by means of an automated script programmed for this purpose.

Four levels of fuel price ranging up to a 150% increase (resulting in CHF 5 per liter) were applied. The extreme scenario incorporated a 50% increase in public transport fares, because higher prices are unrealistic in the short term to midterm. More conservative value ranges were applied for travel times, because large infrastructure improvements are planned on neither the road nor the rail network, and thus travel times should not vary substantially in the short term. In the same vein, to avoid irrelevant planning situations and create unrealistic scenarios, only the status quo or improvements in the number of vehicle changes on the public transport side were considered (that is, connections are not assumed to worsen). The experimental designs for the SP experiments were determined by using the software Ngene (*27*). Every respondent was faced with six mode choice situations, which were displayed after a recapitulative overview of their reported trip.

To clarify the underlying assumptions to the respondents, explanations on how the costs of car and rail trips were computed were displayed for each choice situation. Specifically, the fuel prices applied for the computation of the total car trip cost as well as the assumptions made for the rail fare were detailed. The latter was of utmost importance for owners of Generalabonnement (GA) transit passes. These flat-rate cards, bought once a year, entitle the holder to free use of public transport on the complete Swiss rail network and on local networks. To avoid confusion about why the fares in such cases

TABLE 1 Variable Values Used for Construction of Stated Preference Experiments

Alternative	Attribute	Values
Mode Choice Experiment		
Car	Total travel time	Sum of free-flow and congested travel time
	Free-flow travel time	−10%, ±0%, +10% of current level
	Congested travel time	0%, 10%, 20% of free-flow travel time
	Fuel price	−10%, +50%, +100%, +150% of current level
Rail	Total travel time	Sum of in-vehicle travel time and waiting time
	In-vehicle travel time	−10%, ±0%, +10% of current level
	Waiting time (at transfer)	0, 10, 15 min.
	Number of transfers	0, 1 times
	Fare	−10%, +20%, +50% of current level
Fleet Choice Experiment		
Car	Fixed costs	+20%, +60% of current level
	Fuel price	−10%, +60%, +140% of current level
	Fuel consumption	−25%, −10% of current level
Rail	Fares	−10%, +20%, +50% of current level
	Modal share	10%, 30%, 70% of total yearly mileage

were not set to zero, additional information about how the fare was calculated (cost of the transit pass, CHF 3,100, divided by the total yearly rail mileage) was displayed.

Long-Term Decisions

The long-term SP experiments consisted of six situations in which the characteristics of the respondents' mobility tool fleet were varied along with the distribution of yearly mileage. The attribute levels used for the construction of the experiments are displayed in Table 1. An example situation is displayed in Figure 1. Respondents choose either to keep their current set (under the new fuel and public transport pricing scheme) or switch to an alternative. Here, fuel price is assumed to increase to CHF 3.20 per liter and public transport fares to increase by 50%. The alternative set is constructed as follows. The mileage distribution determined by the experimental design is used to compute total public transport costs for three transit pass settings (no transit pass, half-fare card, and GA), the cheapest of which is chosen and displayed in the questionnaire. The other variables—fuel consumption of the alternate car and its fixed yearly costs (resulting from a distribution of purchase costs over the average lifespan of a car)—result from the experimental design.

Response Rates

The overall response rate was 58.3% (579 respondents), which can be considered as satisfactory considering the questionnaire's complexity and length. The rate matches the experience with comparable studies at IVT, as shown in Figure 2. The ex ante response burden for the different surveys was determined according to the scheme detailed in the work of Axhausen and Weis (*28*). The methodology assigns weighted scores to question types and sums them to calculate the response burden of the survey. The present study fits in the corresponding context for surveys with prior recruitment.

Nontrading Behavior

Nontraders are respondents in SP surveys who, regardless of the alternatives' attributes, always pick the same alternative. A rather high share of respondents (roughly 47%) in the mode choice experiments were nontraders. The share is slightly higher among public transport users than among those who had chosen the car in their reference trip. This may have several reasons, one of which is randomly picking the first or second alternative for every situation to reduce the mental effort of completing the questionnaire. Such behavior evidently biases the outcomes of the statistical analysis of the data, because these effects can lead to decisions that are not based on trade-offs between the alternatives' attributes (*29*). However, especially when respondents have a strong prior commitment to a specific transport mode, apparently illogical choices may well reflect true behavioral response. The most prominent examples of such effects are residential location (where the location type dictates the use of a certain means of transport, for example, the car in regions that are not or are barely accessible by public transport) or transit pass ownership (such as the afore-

Behavior	Current	Alternate
Transit pass	Half-fare card	GA
Price of transit pass	250.-CHF/year	4,500.-CHF/year
Public transport mileage	9,000 km/year (about 30%)	22,000 km/year (about 70%)
Total public transport costs	2,200.-CHF/year	4,500.-CHF/year
Car	Current	New
Fuel consumption	6.0 l/100 km	5.4 l/100 km
Fixed costs	600.-CHF/year	700.- CHF/year
Car mileage	22,000 km/year (about 70%)	9,000 km/year (about 30%)
Fuel costs*	4,200.-CHF/year	1,600.-CHF/year
Total car costs	4,800.-CHF/year	2,300.-CHF/year
Total mobility costs	7,100.-CHF/year	6,800.-CHF/year
	☐ ← Your choice →	☐

* Result from an assumed fuel price of 3.20 CHF/l.

FIGURE 1 Example of fleet choice experiment (original is in German).

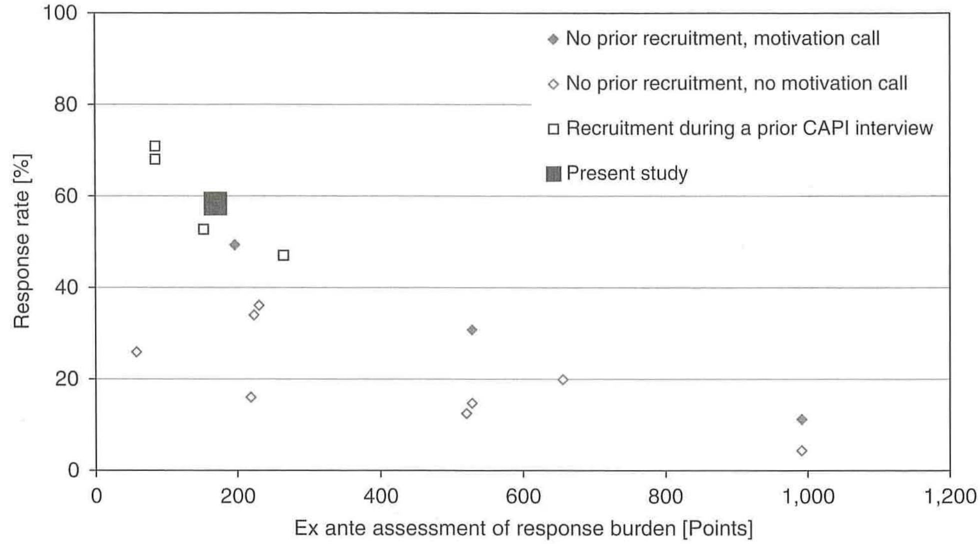


FIGURE 2 Response rate in context of comparable studies.

mentioned GA, which binds customers to the public transport alternative).

Thus, nontrading in itself does not necessarily imply inconsistent responses. Hence, rather than excluding nontraders from the analysis (which would drastically reduce the sample size), special care was taken to model the possible influences on individuals' choice behavior, such as variables describing the trip (distance and purpose), sociodemographic characteristics (especially the availability of mobility tools) of the respondents, and the mode choice in the reference case as an inertia indicator. Obviously, mode choice is determined jointly by all these attributes and not solely by travel times and costs.

In the long-term experiment, only 30% of the respondents were nontraders. This is somewhat surprising because of the slightly more complex nature of the second SP experiment, which as stated above would naturally tend to increase the share of random responses. However, the experiment clearly indicated the cost implications of choosing a specific alternative, thus simplifying the process for the respondents. Individuals appear to be more willing to adapt their behavior in the long than in the short run.

SAMPLE SUMMARY STATISTICS

Table 2 presents key statistics for the respondent sample in comparison with (*a*) the KEP base sample of recruited individuals and (*b*) the 2005 Swiss National Household Travel Survey (*Mikrozensus Verkehrsverhalten* 2005, or MZ '05), which is representative of the Swiss population (*30*).

A bias toward well-educated, rather wealthy respondents can be seen (although the income figures cannot be directly compared to those of the MZ '05, as different measures were used). Single-person households are slightly underrepresented. A high share of respondents own transit passes, as is common for surveys of the described type—users of public transport tend to be more interested in transport policy issues, leading to a higher propensity to participate in the survey. Because only holders of driving licenses were sampled for the study and respondents for which car availability is impossible were excluded, the share of those regularly having a car at their disposal is slightly higher in the KEP than in the population. The share of respondents without a car is higher in the final sample, again indicating an affinity toward public transport users.

The discrepancy of the sociodemographic attributes between the sample and the Swiss population raises the question of the necessity of sample reweighting. As has been stated in the literature (*31*), weighting should be applied in explorative analyses. For the estimation of

TABLE 2 Sample Descriptive Statistics

Variable	Value	Sample	KEP	MZ '05
Gender	Male	48.1	49.4	48.7
	Female	51.9	50.6	51.3
Age (in years)	18–35	17.7	12.3	28.4
	36–50	38.2	39.1	30.4
	51–65	30.5	39.2	23.2
	>65	13.7	9.4	18.1
Education level	Primary or secondary school	6.9	8.0	17.9
	Professional school	47.6	48.3	59.6
	College or university	45.5	43.8	22.5
Number of persons in household	1	15.2	14.3	18.9
	2	41.9	36.4	36.9
	3	13.1	17.1	16.5
	4	17.7	20.0	18.3
	>4	12.1	12.0	9.4
Income (in CHF/ month)	<2,000	16.3	19.0	
	2,000–4,000	17.4	19.3	
	4,000–6,000	28.7	29.2	
	6,000–8,000	19.1	17.0	
	8,000–10,000	10.8	8.7	
	>10,000	7.8	6.8	
Transit pass	None	42.7	51.2	67.2
	Half-fare card	44.8	39.3	26.5
	Generalabonnement	12.5	9.5	6.3
Car availability	Always	77.2	79.8	79.4
	Sometimes	15.5	15.0	15.7
	Never	7.3	5.8	4.9
Car fuel consumption (in l/100 km)	<5	6.5	6.0	
	5–8	69.6	70.2	
	8–12	21.6	21.7	
	12–15	1.9	1.7	
	>15	0.4	0.4	

the discrete choice models described here, weighting is not necessary under the condition that the selectivity variables are included in the models (*31*).

FORMULATION AND ESTIMATION OF DISCRETE CHOICE MODELS

The models for the tactical and strategic decisions make use of discrete choice methodology (*31, 32*), specifically the multinomial logit model. All models were estimated by using the Biogeme software package (*33, 34*).

Mode Choice

Mode Choice Model Formulation

The formulation of the mode choice models follows the form that was introduced in the study by Mackie et al. (*35*) and has been used in several Swiss studies (*3–6*). It adds continuous interactions between variables to the linear utility formulation. As Hess et al. state, the methodology's advantages over an arbitrary segmentation into classes or random parameter models include the computation of deterministic taste heterogeneity and faster estimation times (*5*). The general specification of the utility function is as follows:

$$f(y,x) = \beta_x \cdot \left(\frac{y}{\bar{y}}\right)^{\lambda_{y,x}} \cdot x \qquad (1)$$

where

x = (dis)utility generating variable, such as travel time or cost;
β_x = utility parameter associated with x, to be estimated;
y = variable assumed to interact with x, such as income or trip distance;
$\bar{y}$ = reference value for variable y, such as the sample mean or median; and
$\lambda_{y,x}$ = elasticity of the influence of y on the (dis)utility generated by x, to be estimated.

Traveler sensitivity to attribute x is assumed to vary with the value of attribute y. In the present case, income and trip distance are assumed to influence the disutility generated by travel time and cost. Normalizing y with its mean ensures that the linear parameter indicates the valuation of x at that point (as the interaction term then equals 1). Additionally, it was assumed that the valuation of travel time and costs would differ between trip purposes. The formulation of the final model therefore includes segmentation into the four categories commuting (work or education), shopping, business, and leisure.

Mode Choice Results

Parameter estimates for the final model are displayed in Table 3. A linear utility specification was estimated first, then the nonlinear interactions and the purpose segmentation were gradually added. The results for the purpose-specific nonlinear model (purpose-specific parameters could be found only for the travel time and cost variables) are shown. Parameter values are displayed along with their *t*-statistics

TABLE 3 Parameter Estimates and Model Fit for Mode Choice Model

		Commuting $n = 592$		Shopping $n = 438$		Business $n = 306$		Leisure $n = 1,975$	
Attribute	Parameter	Est.	*t*-Stat.	Est.	*t*-Stat.	Est.	*t*-Stat.	Est.	*t*-Stat.
Car Alternative									
Inertia	β_{car}	1.492	12.13	1.492	12.13	1.492	12.13	1.492	12.13
Congestion	β_{cong}	−1.456	−2.82	−1.456	−2.82	−1.456	−2.82	−1.456	−2.82
Car availability	β_{car_avail}	0.202	2.69	0.202	2.69	0.202	2.69	0.202	2.69
Travel time	$\beta_{tt_{car}}$	−0.029	−1.77	−0.042	−4.82	−0.005	−0.74	−0.031	−7.79
	$\lambda_{dist,tt_{car}}$	−1.909	−2.78	—	—	−2.020	−3.25	−0.551	−5.28
	$\lambda_{inc,tt_{car}}$	1.144	2.08	—	—	3.933	1.49	−0.210	−6.26
Fuel cost	$\beta_{cost_{car}}$	−0.085	3.29	−0.044	−2.61	−0.039	−2.60	−0.038	−7.66
	$\lambda_{dist,cost_{car}}$	—	—	—	—	—	—	−0.261	−2.63
	$\lambda_{inc,cost_{car}}$	−0.985	−3.38	−1.427	−3.50	—	—	−0.184	−1.28
Public Transport Alternative									
Transfers	$\beta_{transfers}$	−0.328	−2.42	−0.328	−2.42	−0.328	−2.42	−0.328	−2.42
Waiting time	$\beta_{waiting\_time}$	−0.020	−1.70	−0.020	−1.70	−0.020	−1.70	−0.020	−1.70
Half-fare card	β_{HTA}	1.301	12.57	1.301	12.57	1.301	12.57	1.301	12.57
GA	β_{GA}	1.891	12.10	1.891	12.10	1.891	12.10	1.891	12.10
Travel time	$\beta_{tt_{public}}$	−0.037	−3.34	−0.020	−3.49	−0.050	−3.25	−0.016	−8.27
	$\lambda_{dist,tt_{public}}$	−1.034	−2.15	—	—	−2.693	−3.61	−0.632	−5.51
	$\lambda_{inc,tt_{public}}$	0.617	1.19	1.210	1.90	—	—	—	—
Fare	$\beta_{cost_{public}}$	−0.046	−2.56	−0.040	−2.54	−0.039	−2.60	−0.047	−6.26
	$\lambda_{dist,cost_{public}}$	—	—	—	—	—	—	−0.261	−2.63
	$\lambda_{inc,cost_{public}}$	−0.985	−3.38	−1.427	−3.50	—	—	−0.184	−1.28
								Adjusted $\rho^2 = 0.305$	

NOTE: n = number of observations, est. = estimate, *t*-Stat. = *t*-statistic, dist = distance, inc = income.

TABLE 4 Population-Weighted WTP Indicators and Demand Elasticities

	Commuting	Shopping	Business	Leisure
WTP Indicator				
Car travel time [CHF/h]	24.4 ± 7.3	55.9 ± 5.9	81.6 ± 61.6	94.4 ± 99.9
Rail travel time [CHF/h]	31.7 ± 5.6	29.1 ± 3.4	185.2 ± 204.2	43.3 ± 5.9
Transfer waiting time [CHF/h]	27.5	27.5	27.5	27.5
Number of transfers [CHF/transfer]	6.4	6.4	6.4	6.4
Travel Time Elasticities				
Car	−0.26	−0.49	−0.13	−0.71
Public transport	−1.29	−1.34	−1.39	−2.37
Price Elasticities				
Car	−0.23	−0.19	−0.46	−0.37
Public transport	−0.30	−0.30	−0.40	−0.56

(absolute values above 1.96 indicate significance of the parameters at the 5% level), as well as general model fit information. Only parameters with *t*-values above 1 were retained in the final model.

The adjusted ρ^2 value of 0.3 indicates a good model fit, and all the included variables are of the expected sign and statistically significant. The nonlinear interaction terms (that is, the λ's from the table) imply that

- With increasing trip distance, travel time and cost sensitivity decrease (negative signs for the parameters), and
- With increasing income, sensitivity to travel time increases (positive sign), whereas sensitivity to cost decreases (negative sign).

Table 4 shows the relevant WTP indicators and demand elasticities for the mode choice model. The WTP indicators are obtained by dividing an attribute's parameter by the cost parameter of the according alternative. As the valuation of the travel time and cost parameters vary with trip distance and income, so do the WTP indicators: WTP decreases with trip distance and increases with income (*1*). The values indicated in the table are weighted to represent distance and income averages for the respective trip purposes. Reweighting was done according to the procedure detailed by Hess et al. (*5*). The 95% confidence interval is indicated along with the mean values. The variation should be considered when the results are used to carry out cost–benefit analyses. For the WTP indicators relating to waiting time and number of transfers, the weighting procedure was not applied, because these attributes were included as linear terms in the utility function.

The value of travel time saving for business purposes has a very large confidence interval. The sample size for its computation being quite small (as few business trips were present in the sample), the values should be regarded with caution. In general, the willingness to pay for saving car travel time tends to be larger than that for public transport. This is in line with expectations, as car travelers are more sensitive to travel time than are users of public transport.

Fleet Choice

Fleet Choice Model Formulation

The utility functions for the fleet choice model were formulated as a linear combination of the attributes of the current and alternate fleets. Additionally, fuel price (in Swiss francs per liter) and the age of the current car were incorporated in the utility function for the alternate fleet, because both variables were expected to directly influence a respondent's decision to purchase a new car. Two inertia variables were included in the model: a constant term for the current alternative (modeling the base utility of that alternative as compared to the other) and a dummy variable indicating whether the transit pass present in the alternative considered was the same as the current one. The model was estimated as a panel mixed logit model; a normally distributed error term was included to model taste heterogeneity between respondents.

Fleet Choice Results

All parameters (Table 5) are of the expected sign and statistically significant. Although inertia is again obviously present (as highlighted by the large constant term), the random parameter associated with the inertia variables is quite large, signaling that there is a subset of travelers who are less inert and are willing to reconsider their choices when prices change. Car costs are perceived more negatively than public transport costs, and fixed costs more negatively than variable costs. The presence of a transit pass, especially the GA, has a positive effect on an alternatives' utility. Reallocating mileage from the car to public transport is valued negatively. The assumption that both fuel price and the age of the current car have a positive effect on the probability of choosing a new fleet is confirmed by the corresponding parameter estimates.

Trade-offs and price elasticities for the fleet choice model are displayed in Table 6. The trade-offs for the respective cost components indicate their relative valuation. That fixed car costs are valued more negatively than fuel costs indicates a lack of willingness to invest in a new, less consumption intense but more expensive vehicle, even with fuel prices rising significantly. On the other hand, the willingness to pay for lower fuel consumption, which is given by the ratio of the parameters for the fixed costs and consumption, amounts to CHF 320 per liter saved on a 100-km distance. Aggregated over the assumed 10-year life of a vehicle, this results in a total of CHF 3,200.

The ratio of the transit pass price and variable public transport cost parameters is even higher. Thus, respondents are willing to pay the costs for a transit pass only under the condition that a sufficient amount in variable costs can be saved. This is indicated by the ratio

TABLE 5 Parameter Estimates and Model Fit for Fleet Choice Model

Attribute	Parameter	Est.	*t*-Stat.
Constant current fleet	c	2.293	6.17
Standard deviation of the error term	σ	1.997	16.18
Current transit pass	β_{current}	0.643	4.98
No transit pass (reference category)	β_{none}	—	—
Half-fare card	β_{HTA}	0.110	0.70
GA	β_{GA}	1.192	2.02
Price of transit pass	$\beta_{\text{season\_ticket\_price}}$	−0.313	−2.03
Variable public transport costs	$\beta_{\text{pt\_var\_costs}}$	−0.038	−0.51
Fixed car costs	$\beta_{\text{car\_fixed\_costs}}$	−0.912	−9.70
Variable car costs (fuel)	$\beta_{\text{car\_var\_costs}}$	−0.242	−6.40
Yearly public transport mileage	$\beta_{\text{pt\_mileage}}$	−0.049	−2.95
Fuel consumption	$\beta_{\text{fuel\_consumption}}$	−0.292	−3.00
Fuel price (effect on alternate fleet)	$\beta_{\text{fuel\_price}}$	0.351	6.76
Age of current car (effect on alternate fleet)	$\beta_{\text{age}_{\text{car}}}$	0.033	1.51
		Adj. $\rho^2 = 0.266$	

NOTE: Est. = estimate, *t*-stat. = *t*-statistic.

TABLE 6 Trade-Offs and Elasticities for Fleet Choice Model

Trade-Off	Unit	Value
$\frac{\beta_{\text{car\_fixed\_costs}}}{\beta_{\text{season\_ticket\_price}}}$	[—]	2.9
$\frac{\beta_{\text{car\_var\_costs}}}{\beta_{\text{pt\_var\_costs}}}$	[—]	6.4
$\frac{\beta_{\text{car\_fixed\_costs}}}{\beta_{\text{car\_var\_costs}}}$	[—]	3.8
$\frac{\beta_{\text{season\_ticket\_price}}}{\beta_{\text{pt\_var\_costs}}}$	[—]	8.2
$\frac{\beta_{\text{season\_ticket\_price}}}{\beta_{\text{car\_var\_costs}}}$	[—]	1.3
$\frac{\beta_{\text{HTA}}}{\beta_{\text{season\_ticket\_price}}}$	[CHF/Halbtax]	−350.1
$\frac{\beta_{\text{GA}}}{\beta_{\text{season\_ticket\_price}}}$	[CHF/GA]	−3,803.4
$\frac{\beta_{\text{GA}}}{\beta_{\text{pt\_mileage}}}$	[km/GA]	−21,929.1
$\frac{\beta_{\text{car\_var\_costs}}}{\beta_{\text{pt\_mileage}}}$	[km/CHF]	4.9
$\frac{\beta_{\text{fuel\_consumption}}}{\beta_{\text{car\_fixed\_costs}}}$	[CHF/(l/100 km)]	319.9

NOTE: GA: Direct price elasticity, −1.18; cross (fuel) price elasticity, 0.05. Half-fare card: direct price elasticity, −0.04; cross (fuel) price elasticity, 0.25.

of the GA parameter and the costs. Compared to a fleet that does not include a transit pass, the utility of one that has a GA is equal at about CHF 3,800 additional costs. This amount constitutes the willingness to pay for a GA, when the corresponding savings in variable costs are provided. The analogous willingness to pay for a half-fare card is CHF 320.

The ratio for the transit pass price and fuel costs parameters is 1.3. This can be interpreted as follows. When replacing the half-fare card by a GA, equal utility is reached when variable car costs are reduced by CHF 3,700 (the difference in transit pass prices multiplied by the aforementioned factor). At the mean fuel consumption (8 liters per 100 km) and a fuel price of CHF 2 per liter, this corresponds to a reallocation of yearly mileage of about 23,000 km. Thus, acquiring a GA is seen as a commitment to public transport that is amortized by traveling more by that mode.

With increasing fuel prices, the propensity of acquiring a GA and reassigning mileage to public transport increases. Assuming fuel price will double and reach CHF 4 per liter, acquiring a GA at the current costs of CHF 3,100 will be profitable when 11,000 km are shifted from car to rail (as the CHF 3,700 savings in variable car costs is then reached).

From the estimated parameters, probabilities for possessing no annual transit pass, a half-fare card, or a GA, were computed under various pricing scenarios. The results are shown in Figure 3. The fuel price of CHF 2 per liter and zero increase in transit pass costs form the base scenario. At the time of writing, about 28.5% of Swiss residents owned a half-fare card, and approximately 4.9% were GA holders. At the current fare setting, fuel price increases would lead to transit pass holder shares of up to 40% in the scenario of CHF 5 per liter. At the same time, when fuel prices do not increase or do so only moderately (as can be realistically assumed after the recent developments), transit pass sales can be increased only if fares are held constant. As shown in Table 6, the direct price elasticity for the GA is at about −1.2, and the cross elasticity is quite low (*1*).

CONCLUSION AND OUTLOOK

The findings in this paper suggest that even if fuel prices again increase dramatically, inertia is present in both mode choice and mobility tool ownership and mileage distribution. Price elasticities computed from the discrete choice models suggest that respondents are more sensitive to public transport price increases than to rising fuel prices. Thus, if the demand for rail services and transit pass sales is to be maintained, no wide margin exists for increasing public transport fares. Travel times appear to be of greater importance in determining mode choice. These findings are interesting in view of the improvements to the Swiss rail infrastructure that are to be made in the midterm future, such as the opening of the Gotthard Base tunnel. Effects of such improvements on the share of annual transit pass holders cannot be ignored and should be thoroughly investigated in the future, as they hold a significant potential for modal shifts if the fares are held at the current level.

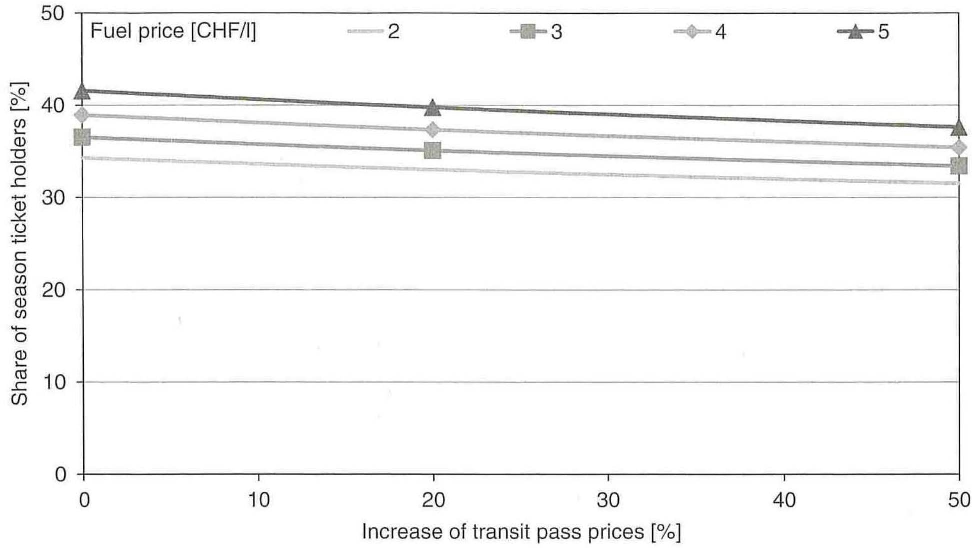


FIGURE 3 Forecast of share of transit pass owners under various price settings.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the Swiss Federal Railway Company and the advice of Tobias Ebinger and Johannes Lieberherr. Furthermore, we would like to thank Urban Schwegler of the Link Institut, Zurich, for assistance in the recruitment process and prompt data delivery throughout the survey period. Alexander Erath contributed programming efforts to the script used to obtain rail travel times, and John Rose provided the Ngene software for the design of the SP experiments.

REFERENCES

1. Weis, C., and K. W. Axhausen. *Benzinpreis und Bahnnutzung.* IVT, ETH Zürich, 2009.
2. Vrtic, M., and P. Fröhlich. Was Beeinflusst die Wahl der Verkehrsmittel? *Der Nahverkehr,* Vol. 24, No. 4, 2006, pp. 52–57.
3. Axhausen, K. W., S. Hess, A. Koenig, G. Abay, J. J. Bates, and M. Bierlaire. Income and Distance Elasticities of Values of Travel Time Savings: New Swiss Results. *Transport Policy,* Vol. 15, No. 3, 2008, pp. 173–185.
4. Axhausen, K. W., S. Hess, A. Koenig, J. Bates, and M. Bierlaire. State-of-the-Art Estimates of Swiss Value of Travel Time Savings. Presented at 86th Annual Meeting of the Transportation Research Board, Washington D.C., 2007.
5. Hess, S., A. Erath, and K. W. Axhausen. Estimated Value of Savings in Travel Time in Switzerland: Analysis of Pooled Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2082,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 43–55.
6. *Zeitkosten im Personenverkehr: Kosten-Nutzen-Analysen im Strassenverkehr.* Swiss Association of Road and Transportation Experts (VSS), Zürich, 2007.
7. Fang, H. F. A Discrete-Continuous Model of Households' Vehicle Choice and Usage, with an Application to the Effects of Residential Density. *Transportation Research B,* Vol. 42, No. 9, 2008, pp. 736–758.
8. De Jong, G., J. Fox, A. Daly, M. Pieters, and R. Smit. Comparison of Car Ownership Models. *Transport Reviews,* Vol. 24, No. 4, 2004, pp. 379–408.
9. Baranzini, A., D. Neto, and S. Weber. *Elasticité-prix de la Demande d'Essence en Suisse.* CRAG, HEC Geneva, and LEA, University of Geneva, 2009.
10. Page, M., G. Whelan, and A. Daly. Modelling the Factors Which Influence New Car Purchasing. Presented at European Transport Conference, Cambridge, United Kingdom, 2000.
11. Scott, D. M., and K. W. Axhausen. Household Mobility Tool Ownership: Modeling Interactions Between Cars and Season Tickets. *Transportation,* Vol. 33, No. 4, 2006, pp. 311–328.
12. Vrtic, M., N. Schuessler, A. Erath, and K. W. Axhausen. Route, Mode, and Departure Time Choice Behaviour in the Presence of Mobility Pricing. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
13. Vrtic, M., N. Schuessler, A. Erath, and K. W. Axhausen. Design Elements of Road Pricing Schemes and Their Acceptability. Presented at 86th Annual Meeting of the Transportation Research Board, Washington, D.C., 2007.
14. Axhausen, K. W., A. Simma, and T. Golob. Pre-commitment and Usage: Cars, Season-Tickets and Travel. *European Research in Regional Science,* Vol. 11, 2001, pp. 101–110.
15. Chatterjee, K., and G. Lyons. Coping with a Crisis: Lessons from the Events of a Year Ago. Presented at European Transport Conference, Cambridge, United Kingdom, 2001.
16. Polak, J. W., R. B. Noland, M. G. H. Bell, N. Thorpe, and D. Wofinden. What Happens when the Pumps Run Dry? Experience from the 2000 Fuel Protest and Its Policy Implications. Presented at European Transport Conference, Cambridge, United Kingdom, 2001.
17. Noland, R. B., J. W. Polak, and M. G. H. Bell. Expected Disruption and Flexibility of Vehicle Users to Fuel Shortages. Presented at 81st Annual Meeting of the Transportation Research Board, Washington, D.C., 2002.
18. *Kontinuierliche Erhebung Personenverkehr.* Swiss Federal Railways, Bern, Switzerland, 2001.
19. Louvière, J. J., D. A. Hensher, and J. Swait. *Stated Choice Methods: Analysis and Applications in Marketing, Transportation and Environmental Valuation.* Cambridge University Press, Cambridge, United Kingdom, 2000.
20. Algers, S., J. Lindqvist, J. Dillen, and S. Wildert. The National Swedish Value of Time Study. Presented at PTRC European Transportation Forum, Warwick, United Kingdom, 1995.
21. Kurri, J., and M. Pursula. *Finnish Preliminary Value of Time Studies.* Swedish Road and Transport Research Institute and Swedish Transport and Communications Research Board, Linköping, Sweden, 1995.
22. Ramjerdi, F., L. Rand, I. Saestermo, and K. Saelensminde. *The Norwegian Value of Time Study.* Norwegian Centre for Transport Research, Oslo, Norway, 1997.

23. Jovivic, G., and C. O. Hansen. A Passenger Travel Demand Model for Copenhagen. *Transportation Research A,* Vol. 37, No. 4, 2003, pp. 333–349.
24. Vrtic, M., and K. W. Axhausen. Impact of Tilting Trains in Switzerland: Route Choice Model of Regional and Long-Distance Public Transport Trips. Presented at 82nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.
25. Vrtic, M., and K. W. Axhausen. Forecast Based on Different Data Types: A Before and After Study. Presented at 10th World Conference on Transport Research, Istanbul, Turkey, 2004.
26. Arendt, M., and M. Vrtic. Nationales Personenverkehrsmodell des Bundes als verkehrsplanerische Grundlage. In *Jahrbuch 2005/2006 Schweizerische Verkehrswirtschaft* (T. Bieger, C. Lässer, and R. Maggi, eds.), SVWG, St. Gallen, Switzerland, 2006, pp. 145–166.
27. Rose, J. M., M. C. J. Bliemer, D. A. Hensher, and A. T. Collins. Designing Efficient Stated Choice Experiments in the Presence of Reference Alternatives. *Transportation Research B,* Vol. 42, No. 4, 2008, pp. 395–406.
28. Axhausen, K. W., and C. Weis. Predicting Response Rate: A Natural Experiment. *Arbeitsberichte Verkehrs- und Raumplanung,* No. 549, Institute for Transport Planning and Systems, Swiss Federal Institute for Technology, Zurich, 2009.
29. Van de Kaa, E. J. Assessment of the Value of Travel Time from Stated Choice Surveys: The Impact of Lexicographic Answering. Presented at European Transport Conference, Strasbourg, France, 2006.
30. Swiss Federal Statistical Office and Swiss Federal Office of Spatial Development. Mobilität in der Schweiz: Ergebnisse des Mikrozensus 2005 zum Verkehrsverhalten. Swiss Federal Statistical Office, Bern, Switzerland, 2007.
31. Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Mass., 1985.
32. Train, K. E. *Discrete Choice Models with Simulation.* Cambridge University Press, Cambridge, United Kingdom, 2003.
33. Bierlaire, M. BIOGEME: A Free Package for the Estimation of Discrete Choice Models. Presented at 3rd Swiss Transport Research Conference, Monte Verità, Switzerland, 2003.
34. Bierlaire, M. *An Introduction to BIOGEME 1.8.* Transport and Mobility Laboratory, École Polytechnique Fédérale de Lausanne. http://biogeme.epfl.ch. Accessed July 3, 2009.
35. Mackie, P., M. Wardman, A. S. Fowkes, G. Whelan, J. Nellthorp, and J. J. Bates. *Values of Travel Time Savings in the UK.* Institute for Transport Studies, University of Leeds, Leeds, United Kingdom, 2003.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Exploring Sense-of-Place Attitudes as Indicators of Travel Behavior

Kate Deutsch and Konstadinos Goulias

Researchers in travel behavior have explored attitudes as important determinants since the 1980s and are now broadening the use of attitudinal data in models to explain behavior. Much of the application, however, has focused on attitudes about entities such as lifestyles and attributes of different transport modes. Little attention has been given to attitudes related to places and attributes of human–place interaction. There has been much theorizing in the past 40 years about this human–place interaction, which has been formalized into a multivoiced theory of sense of place. Quantification of this theory and implementation of this attitudinal information into travel behavior modeling are discussed. A survey was conducted in Santa Barbara, California, to measure sense of place and to estimate models of travel behavior. Patrons of two outdoor shopping malls in Santa Barbara were interviewed about place attitudes, sociodemographic information, and details regarding their activity and travel on that day. Regression models were then used with their data to examine differences between the two study locations, the use of mode for arrival, and the timing of the activity, showing the value of the sense-of-place construct as a determinant of travel behavior.

Recent advances in modeling and simulation of travel behavior have greatly increased the ability to successfully understand, analyze, and predict human behavior as it pertains to travel. Introduction of the activity-based approach has provided a new foundation for modeling practices (*1*). In conjunction with implementation of the activity-based approach, the field is recognizing and responding to the necessity to integrate land use patterns and transportation into an integrated model that account for interactions between the two (*2*). These two advances, along with several additional advancements to modeling techniques, have increased the accuracy and reliability of prediction and simulation efforts. Improvements of traditional discrete choice methods include latent variables and classes (*3, 4*), enriching traditional models with personality indicators (*5–8*), social networking information (*9*), and activity scheduling and planning intentions (*10–14*). Attitudinal variables have proved to be valuable in increasing the fit and explanatory power of regression models. For instance, Kitamura et al. found that attitudes are more strongly associated with travel behavior than land use characteristics (*15*). Some of these studies have focused on the choices individuals make about vehicle type (*16*), mode choice with respect to attitudes about mode performance (*17, 18*), and mode choice with respect to attitudes about urban design and practicality of modes (*19*) and urban design as well as environment and time use attitudes (*15*). Although attitudinal determinants have been proved to be advantageous to modeling travel behavior, their use has been limited in practice. Moreover, attitudes regarding places and respondent views toward specific places, to the authors' knowledge, have not been explored and applied to travel behavior.

At the foundation of attitudes about places is sense of place, which is defined by Tuan as a person's "affective ties with the material environment" (*20*). Tuan, like many others in the field of geography, claims that sense of place is a phenomenological process and must be treated as a highly individualized experience and is therefore a fuzzy concept and difficult to quantify. However, others claimed that sense of place can and should be quantified and applied to various research endeavors (*21, 22*). Work in several fields has collectively improved the theoretical framework of sense of place and related concepts. Places, as theorized by Canter, represent "a confluence of cognitions, emotions and actions organized around human agency" (*23*). Work conducted by Jorgensen and Stedman in quantifying sense of place was based on the theory set forth by Canter and attributed to each of the three commonly cited concepts a three-part attitudinal structure to represent cognitive, affective, and conative processes. These three concepts are place identity, which is "a person's identity with relation to the physical environment" (*24*), place attachment, which is defined as "the positive bond that develops between a person and their environment" (*25*), and place dependence, which is defined as the "perceived strength of association between a person and a place" (*26*). Jorgensen and Stedman juxtapose these concepts with the three components of attitude previously mentioned, place attachment being matched with the affective or emotional portion, place identity with the cognitive portion as it relates to the dimensions that build one's sense of self, and place dependence equating with the conative domain, that is, dependence for a place is expressed in actions at that location (*27*). In addition to this work, several researchers have quantified place concepts and applied them to education (*28*), psychology (*29*), forestry (*30*), architecture (*31, 32*), computer science and geographical information science (*33*), sociology (*34*), and geography (*35*). Most recently, applications to travel behavior have occurred that use sense-of-place theory (*36*) or parts of the theory (*37*). The remainder of this paper discusses in detail the data collection and analysis of sense of place as it relates to travel behavior in Santa Barbara, California.

DATA COLLECTION AND SURVEY PROCESS

Two shopping malls in Santa Barbara were chosen as study sites for this research. Both are outdoor malls and are located minutes off US-101 (a north–south high-level facility connecting Los Angeles to San Francisco, California). Paseo Nuevo mall is situated in the

Department of Geography, University of California, Santa Barbara, 1832 Ellison Hall, Santa Barbara, CA 93106. Corresponding author: K. Deutsch, deutsch@geog.ucsb.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 95–102.
DOI: 10.3141/2157-12

TABLE 1 Sample Descriptive Statistics

	n	Age (mean, years)	Gender (% female)	SB County Residents (%)	Household Income (median) ($)	Household with One or More Persons Younger than 18 Years (%)
Paseo Nuevo	320	34.75	41.6	69.7	55,000–74,999	20.3
La Cumbre	503	39.69	64.8	80.1	55,000–74,999	27.8
Total	823	37.7	55.8	76.1	55,000–74,999	25.2

NOTE: SB = Santa Barbara.

middle of downtown Santa Barbara and abuts State Street, a well-known and common tourist destination. Parking structures are located in several areas surrounding the mall, and there is limited street parking. Architecture in the downtown area of Santa Barbara is mostly of Spanish influence and common characteristics are carried throughout the city blocks. La Cumbre is located on upper State Street in a more-residential area. La Cumbre is surrounded by surface parking lots and has a more-traditional suburban mall appearance. The architectural style of upper State Street is not as unified as that of the lower State Street area. A more complete description of the study areas and the survey process is available elsewhere (*36, 38*).

The sample of this study was recruited through an intercept survey process. Patrons were surveyed at each shopping mall during 5 days of a 1-week period at each location. The resulting sample size of this survey was 823 respondents. The resulting sample is 55.8% female having an average age of 37.7 years. Additionally, 76.1% of the respondents were residents of Santa Barbara, 3.6% were residents of Ventura County (immediately south of Santa Barbara County), 7.7% from neighboring counties such as Los Angeles and San Luis Obispo, and 4.6% from other counties in California. The remaining 8% of the respondents were from out of state or another country or had multiple residences. Statistics for each mall are given in Table 1.

Broken into disaggregate samples at each mall, the resulting sample size was 320 for Paseo Nuevo and 503 for La Cumbre. The Paseo Nuevo sample had an average age of 34.75 years, and 41.6% of the respondents were female. Of the sample, 69.7% were residents of Santa Barbara County, and 20.3% came from a household with one or more persons under the age of 18. The La Cumbre sample had an average age of 39.69 years, and 64.8% of the respondents were female. Of the sample, 80.1% were residents of Santa Barbara County, and 27.8% came from a household with one or more persons under the age of 18.

Following the work of Jorgensen and Stedman (*27*), sense of place is treated as a set of attitudinal components. In addition to place dependence, place identity, and place attachment as discussed earlier, place satisfaction was also measured. Place satisfaction is defined as "a person's level of satisfaction with the services, environment and needs provided for by a specific place" (*39*). Attitudinal measures of aesthetic quality, atmosphere, and cultural and social associations were included in the questionnaire. A list of concepts related to sense of place, from which questions were developed, is provided in Table 2.

SENSE OF PLACE AS DETERMINANT OF TRAVEL BEHAVIOR

To determine the benefit of sense-of-place indicators in modeling of travel behavior, several models were constructed by using a nested specification. The nesting and comparison of several models allows for analysis of the cost (of using more degrees of freedom) versus the benefit (improvement of fit) with the addition of place indicators while considering if the models thus derived are reasonable and in agreement with past findings. Regression models were created to investigate differences between respondents at each location, the modal split of the observed data, and characteristics of the trip, such as timing of the activity. It is important to note that these models do not attempt to analyze the choice of the individual (either the mode or the location), as these models do not consider the entire choice set explicitly nor all the attributes of each alternative (such as cost and distance). Rather, these models explore the addition of sense-of-place indicators in measurement of differences within the sample of a selection of travel behavior attributes.

To examine the impact of sense of place on travel behavior, models were built with location as the dependent variable. A location dummy variable was used as the dependent variable (1 for La Cumbre, 0 for Paseo Nuevo), and a binary Logit model was used to explain the latent propensity of traveling to either mall. This analysis provides insight into the differences between people at each location as well as their attitudes toward the place and its attractiveness as an activity center. Comparison among different models in the nested specification is used to examine the impact of the addition of sense-of-place indicators to a travel behavior model containing only sociodemographic information. The initial model containing only sociodemographic information is given in Table 3. Results for indicators with significance at a .05 level or better are shown in black. Indicators that were bordering on significance (better than significance at .1) are shown in gray.

TABLE 2 Sense-of-Place Concepts and Related Survey Topics

Sense-of-Place Concept	Related Topics
Place attachment	Ability or likelihood of relaxing, happiness due to place, importance of existence, level of importance compared to other places
Place dependence	Needs met, diversity, underlying existence of reasons for trip
Place identity	Identification with atmosphere, place as a reflection of the individual, level of freedom to be self
Place satisfaction	Satisfaction with food, products, parking, level of service, entertainment, crowd size
Aesthetics	Views on architecture, beauty of place, balance of decorative and functional attributes, artistic value, peaceful and relaxing atmosphere
Social and cultural	Social atmosphere, reflects culture of Santa Barbara, risk of unpleasant encounters, level of crowdedness, amount of activity, safety of walking around, family- and kid-friendly, level of friendliness of people

TABLE 3 Binary Logit Location Model

	Sociodemographic					Sociodemographic–Sense of Place				
Model Variable	B	SE	*t*-Stat.	Sig.	ExpB	B	SE	*t*-Stat.	Sig.	ExpB
Constant	.300	.298	1.008	.313	1.350	1.318	.674	1.956	.050	3.737
MALE (1 if male; 0 otherwise)	–.874	.159	5.505	.000	.417	–.682	.255	2.672	.008	.505
AGE18_24 (1 if age 18–24; 0 otherwise)	–1.177	.399	2.951	.003	.308					
AGE25_29 (1 if age 25–29; 0 otherwise)	–.704	.424	1.662	.097	.495	–1.010	.632	1.599	.110	.364
AGE30_39 (1 if age 30–39; 0 otherwise)	–.882	.387	2.278	.023	.414	–1.070	.588	1.818	.069	.343
AGE40_65 (1 if age 40–65; 0 otherwise)	–.666	.320	2.086	.037	.514					
SBRES (1 if home in Santa Barbara County; 0 otherwise)	.618	.183	3.382	.001	1.856					
INCHIGH (1 if household income higher than $100,000; 0 otherwise)						–1.130	.366	3.091	.002	.323
MARRIED_DP (1 if married or domestic partner; 0 otherwise)	.586	.269	2.182	.029	1.797					
EMPSTUD (1 if student; 0 otherwise)	.532	.330	1.613	.107	1.703					
EMPPART (1 if employed part-time; 0 otherwise)	.527	.335	1.575	.115	1.694	1.059	.520	2.036	.042	2.882
ONEKID (1 if one kid in the household; 0 otherwise)	.449	.262	1.716	.086	1.567	.957	.432	2.213	.027	2.603
KIDS2_MORE (1 if two or more kids in the household; 0 otherwise)	.368	.259	1.418	.156	1.444	1.001	.401	2.498	.012	2.722
CARS2 (1 if two cars in household; 0 otherwise)	.819	.253	3.233	.001	2.268					
"I am satisfied with the products" (1 if strongly dissatisfied, dissatisfied, or slightly dissatisfied; 0 otherwise)						.797	.376	4.485	.034	2.219
"I am satisfied with the products offered" (1 if neutral; 0 otherwise)						1.002	.322	9.692	.002	2.725
"I am satisfied with the parking" (1 if strongly dissatisfied, dissatisfied, or slightly dissatisfied; 0 otherwise)						–2.797	.400	48.934	.000	.061
"I am satisfied with the parking" (1 if neutral; 0 otherwise)						–1.746	.445	15.416	.000	.174
"I am satisfied with the service" (1 if slightly agree or agree; 0 otherwise)						–1.012	.417	5.890	.015	.364
"I am satisfied with the entertainment" (1 if strongly dissatisfied, dissatisfied, or slightly dissatisfied; 0 otherwise)						2.043	.477	18.317	.000	7.715
"I am satisfied with the entertainment" (1 if neutral; 0 otherwise)						1.565	.449	12.148	.000	4.781
"[Loc] has beautiful architecture" (1 if strongly dissatisfied, dissatisfied, slightly dissatisfied, or neutral; 0 otherwise)						1.463	.414	12.509	.000	4.320
"[Loc] has a definite social atmosphere" (1 if agree; 0 otherwise)						–1.669	.282	34.907	.000	.188
"[Loc] makes me feel happy" (1 if neutral; 0 otherwise)						.890	.303	8.659	.003	2.436
"[Loc] is kid friendly" (1 if strongly dissatisfied, dissatisfied, or slightly dissatisfied; 0 otherwise)						–1.475	.492	9.004	.003	.229
"[Loc] is kid friendly" (1 if neutral or slightly agree; 0 otherwise)						–.677	.280	5.861	.015	.508
"[Loc] does not involve the risk of unpleasant encounters" (1 if strongly disagree, disagree, or slightly disagree; 0 otherwise)[a]						–.981	.339	8.371	.004	.375
"[Loc] is not overcrowded" (1 if agree; 0 otherwise)[a]						.851	.279	9.300	.002	2.343
"[Loc] is not overcrowded"(1 if strongly agree; 0 otherwise)[a]						1.754	.426	16.997	.000	5.780
"[Loc] makes me feel comfortable because I identify with the atmosphere" (1 if neutral; 0 otherwise)						.802	.306	6.884	.009	2.230
$n = 823$, X^2 improvement (Model 1 to 2) = 503.759, change in df = 21	–2 LL = 986.336, df = 18, $X^2 = 113.551$					–2 LL = 482.577, df = 39, $X^2 = 617.310$				

NOTE: Only significant variables reported; loc = location.
[a]Indicates question was recoded in inverse.

Results indicate that age has significance in the differences between locations, with younger age groups being less likely to frequent La Cumbre (i.e., the suburban-like mall). Perhaps the most noteworthy aspect of age is shown in the result for the college-age group (ages 18 to 24), which indicates a considerably larger aversion to traveling to La Cumbre, with greater significance than the other age groups. The results, however, indicate that age does not have a significant contribution to the overall log odds (column labeled Exp B) of one choosing La Cumbre over Paseo Nuevo. Gender also has significance, with a negative coefficient for males, indicating that males are less likely to travel to La Cumbre than to Paseo Nuevo. This indicator again contributes less to the overall probability of frequenting La Cumbre. Residency in Santa Barbara is also significant for location, with a higher contribution (1.856), which can most easily be explained by the tourist nature of Paseo Nuevo and its proximity to numerous other tourist attractions. People who are married or who have domestic partners also have a higher propensity to travel to La Cumbre, as do people who have one child in the household and people with two cars. All these indicators have a higher contribution to the probability of a person choosing La Cumbre as a destination.

To examine the use of sense-of-place indicators to this model, more indicators were added to the initial model. The resulting model is shown in Table 3. These variables indicate the respondent's views regarding aspects of sense of place with respect to the mall that he or she was visiting at the time of the survey. Appropriate dummy indicators were created for sense-of-place questions by using the distributions of each question. Indicators are again shown for a .05 significance level or better and in gray for a 0.1 significance level or better.

The model resulting from the addition of attitudinal indicators presents many interesting results, both in analysis of the resulting significant attitudinal variables and in exploration of both the attrition and the addition of some key sociodemographic information. All age-group dummy variables became insignificant with the addition of attitudinal information (at a .05 level) with the exception of the 30-to-39 age group. This may because of a strong correlation between sense-of-place attitudinal variables and age; for example, specific attitudes are formed at specific stages in life and those are captured by the age variables in one model and the attitudinal variables in the other. In addition to the loss of age indicators, Santa Barbara residency also fell out of the resultant regression model. This is another aspect of sense of place that is worth investigating. People who are residents of Santa Barbara might have similar opinions and attitudes, possibly shifting the explanatory power attributed to the residency indicator to several sense-of-place indicators. A further investigation into the relationship between sense of place and level of familiarity or experience with the location is needed to determine whether this explanation is accurate. Indicators for student status and number of cars were among those that were no longer significant in the model. With the addition of sense-of-place indicators, variables indicating children in the household increased in contribution, and in the case of the indicator for two or more children, increased in significance. Both of these indicators lead to a positive affinity toward traveling to La Cumbre. High-income households also became significant with the addition of attitudes, with a negative coefficient (indicating a lower likelihood to travel to La Cumbre), but a relatively low overall contribution to the probability.

Several attitudinal variables are significant in predicting location, and indicators measuring the respondent's satisfaction with different aspects of each place are also significant. People at Paseo Nuevo, for instance, have a higher tendency to be unsatisfied to some degree or to have no opinion about the products offered at the mall. This could suggest that there are other factors that attract visitors to Paseo Nuevo. Variables indicating product satisfaction have a similar contribution to the overall probability as several sociodemographic indicators. La Cumbre patrons were more likely to be dissatisfied to some degree with entertainment options, with a very large contribution to the overall probability of one traveling to La Cumbre as indicated by the log odds (7.715 for disagreeing). These results suggest lack of entertainment options besides shopping. This is justified by several qualitative comments offered in an open-ended section of the survey:

- "Paseo Nuevo is more of a hang out, La Cumbre is for specific shopping."
- "Paseo has a lot more entertaining and more enjoyable to come to than La Cumbre."
- "La Cumbre allows for wandering and shopping. Paseo Nuevo seems fun, but because of parking limits, I feel rushed."
- "No movies at La Cumbre. No live theater either."
- "La Cumbre is more for locals and people who just want to shop for a specific thing, and they can just get in and get out. They don't have to deal with traffic or parking. Paseo Nuevo is more for going out downtown for the night so it's more for an occasion."

Several other aspects of sense of place are shown to be significant. Differences in attitudes about the number of people at each location are shown to be significant. Those who disagree with the statement, "[mall of patronage] is too overcrowded" are more likely to be at La Cumbre. The distributions of responses at each location indicate a definite disagreement with the statement at La Cumbre but a bimodal distribution at Paseo Nuevo of either disagreement or indecisive attitudes. Those who agree with the statement that "[Paseo Nuevo/La Cumbre] involves the risk of unpleasant encounters" (or, again, the inverse) are more likely to travel to Paseo Nuevo than to La Cumbre. Both aspects of sense of place could be attributed to the locality of the mall, Paseo Nuevo being surrounded by the downtown area and La Cumbre being somewhat isolated and surrounded by large parking lots, and a more suburban design to the neighborhood. In addition, people who disagree that where they are is kid friendly are less likely to travel to La Cumbre. It can be determined through these indicators that respondents at Paseo Nuevo view the location to which they traveled as a social environment with good entertainment but as more risky and perhaps less car friendly, compared to the views of those who traveled to La Cumbre.

To determine the appropriateness of the addition of attitudinal variables, a chi-squared analysis was conducted. Table 3 provides model fit results for each of the models previously discussed, as well as a difference in chi-squared statistics and degrees of freedom. To test the influence of the additional sense-of-place indicators, a null hypothesis assuming no influence of attitudinal indicators was established. This null hypothesis would statistically be true until the threshold value of 46.80 for a change in chi-squared given the difference of 21 degrees of freedom. The change in chi-squared of 503.759 provides statistical grounds to reject the null hypothesis.

To further test the benefit of adding sense-of-place indicators, an examination of mode was conducted. For this analysis, a stepwise nesting procedure was again utilized. A model of modal split using only sociodemographic information was estimated first and is shown in Table 4. Model results indicate that Santa Barbara residency, age, gender, marital status (at a .1 significance level), and car ownership all have an impact on walking. Being a resident of Santa Barbara has a negative impact on one's likelihood to walk, as does being female;

TABLE 4 Modal Split Model

	Sociodemographic					Sociodemographic/Sense of Place				
Model of Transport	B	SE	*t*-Stat.	Sig.	ExpB	B	SE	*t*-Stat.	Sig.	ExpB
Walk										
Intercept	.383	.432	0.888	.375		.486	.656	0.741	.459	
FEMALE (1 if female; 0 if male)	−.630	.318	3.926	.048	.532	−.569	.334	1.702	.089	.566
AGE40_65 (1 if age 40–65; 0 otherwise)	.887	.389	5.203	.023	2.428	.848	.404	2.100	.036	2.335
SBRES (1 if residence in Santa Barbara County; 0 otherwise)	−1.078	.440	5.996	.014	.340	−1.026	.459	2.234	.026	.358
MARRIED_DP (1 if married or domestic partnered; 0 otherwise)	.752	.444	1.692	.091	2.120					
CARS1 (1 if one car in household; 0 otherwise)	1.416	.408	12.047	.001	4.120	1.376	.426	3.233	.001	3.960
CARS2 (1 if two cars in household; 0 otherwise)	1.086	.480	5.112	.024	2.963	1.052	.496	2.119	.034	2.862
CARS3UP (1 if three or more cars in household; 0 otherwise)	1.176	.463	2.541	.011	3.241	1.137	.477	2.385	.017	3.117
"I am satisfied with the people" (1 if agree or strongly agree; 0 otherwise)						1.084	.360	9.043	.003	2.956
"[Loc] has a peaceful and relaxing atmosphere" (1 if slightly agree or agree; 0 otherwise)						−.588	.35	2.802	.094	.556
"[Loc] has a definite social atmosphere" (1 if agree; 0 otherwise)						−.929	.373	6.198	.013	.395
"I am not afraid to walk around at [Loc]" (1 if slightly disagree, neutral or slightly agree; 0 otherwise)[a]						.823	.497	2.744	.098	2.277
Car										
Intercept	1.294	.380	3.404	.001		1.538	.567	2.713	.007	
SBRES (1 if residence in Santa Barbara County; 0 otherwise)	−1.235	.398	3.106	.002	.291	−1.445	.416	3.476	.001	.236
MARRIED_DP (1 if married or domestic partnered; 0 otherwise)	1.326	.389	3.410	.001	3.764	1.109	.399	2.775	.006	3.030
KIDS2_MORE (1 if two or more kids in the household; 0 otherwise)	−1.008	.396	2.544	.011	.365	−1.002	.411	2.435	.015	.367
CARS1 (1 if one car in household; 0 otherwise)	1.631	.340	4.792	.000	5.111	1.549	.358	4.324	.000	4.707
CARS2 (1 if two cars in household; 0 otherwise)	1.934	.400	4.836	.000	6.914	1.734	.416	4.165	.000	5.664
CARS3UP (1 if three or more cars in household; 0 otherwise)	1.990	.379	5.256	.000	7.314	2.019	.397	5.092	.000	7.532
"I am satisfied with the parking" (1 if slightly agree or agree; 0 otherwise)						.951	.333	8.177	.004	2.588
"I am satisfied with the parking" (1 if strongly agree; 0 otherwise)						.846	.352	5.766	.016	2.330
"[Loc] has a definite social atmosphere" (1 if agree; 0 otherwise)						−1.349	.316	18.231	.000	.259
"[Loc] makes me feel happy" (1 if neutral; 0 otherwise)						.624	.366	2.912	.088	1.867
Sample size = 823, X^2 improvement (Model 1 to 2) = 84.725, change in df = 18	−2LL(0) = 470.658, −2 LL(β)= 339.427, df = 16, X^2 = 131.256					−2LL(0) = 1.099E3, −2 LL(β) = 482.577, df = 34, X^2 = 215.981				

NOTE: Only significant variables reported; reference category = "other"; loc = location.
[a]Indicates question was recoded in inverse.

however, car ownership, being between the ages of 40 and 65, and being married or having a domestic partner have positive impacts. Having two cars has the least positive impact on an individual's likelihood to walk within the car ownership indicators. All three indicators for car (with the reference indicator being zero cars in the household) have a relatively high log odds ratio, meaning that they contribute significantly to the overall probability of walking. Additionally, indicators for being married and for being between 40 and 65 contribute significantly to the overall likelihood of a person walking. Car as a mode also had several significant indicators. Indicators with highest contributions continued to be car-ownership dummy indicators. Comparison of log odds show that household car ownership contributes more significantly to the likelihood of arriving by car, compared to foot. Being married also positively and greatly contributes to the overall probability to use car over other modes of transportation. Indicators for two or more children in the household and Santa Barbara residency negatively affect the likelihood for car use. The negative contribution of Santa Barbara residency might be

due to the number of students who ride the bus or downtown residents who use alternative modes to travel to Paseo Nuevo. The resulting negative contribution of those who come from households with more than two children is an interesting result. Perhaps one of the members of the house on lunch break or running errands from work downtown walked to the mall location, or people were on vacation and stayed in a hotel close to the Paseo Nuevo area. Of the 26 respondents who had more than two children in the household and did not drive, 17 visited Paseo Nuevo and nine visited La Cumbre. About half these 26 people were working the day of the survey. Of those who were not working (15 people), seven resided in a county other than Santa Barbara or Ventura.

To examine the impact of sense-of-place indicators on the models, a new model was estimated through use of the same socio-demographic indicators as the initial model with added sense-of-place variables. Results of this model are shown in Table 3. Consistently, car-ownership indicators remained significant for all modes and had large contributions to the likelihoods as indicated by the log odds ratios. Santa Barbara residency remained significant for all modes, probably either indicative that a higher percentage of groups such as students use the bus or because of the lack of tourists who use alternatives other than walking. Several other indicators that remained significant, as discussed previously, are reported in the table. The use of attitudinal indicators provides insight and value for the resulting model and significant attitudes that can predict mode. Satisfaction with the number of people has a positive impact on the likelihood of one walking to the destination. Disagreement with the statement about feeling safe in walking around the location also has a negative impact on walking. Perhaps this is because walkers absorb both positive and negative aspects of the atmosphere. Another possible explanation is that people who work in the area walk for errands or to lunch and could have a less idyllic or leisure-oriented view of the place. Likewise, agreeing with that there is a social or a peaceful and relaxing atmosphere causes one to be less likely to walk to the location. This result indicates that walking as a mode is not necessarily linked to leisure or social trips. Examination of indicators significant for explaining car usage shows, as expected, that positive attitudes toward satisfaction with parking contributes positively to the probability of one traveling by auto. Agreeing with the social nature of the atmosphere contributes negatively toward car use, and having an indifferent opinion about whether the mall makes the individual happy contributes positively to use of an auto (although only at a .1 significance level). Several indicators of sense of place proved to be significant, but to test the value of the additional data to the model a chi-squared comparison was used. Table 4 provides the goodness of fit statistics for the aforementioned models as well as the resulting improvement in chi-squared by adding sense-of-place variables to the model specification. With a threshold of change in chi-squared of 42.31 (from a chi-squared table) for a change of 18 degrees of freedom, one can see that sense-of-place indicators enhance the explanatory power of the model. The difference in chi-squared compared to the threshold value for the modal split model (84.725 versus threshold of 42.31) was not as large as that of the location model (503.759 versus threshold of 46.80). Indicators describing the attributes of the location (distance from last activity, accessibility by mode, etc.), the mode (cost of travel by mode and alternatives, etc.), or the individual's activity pattern (preceding or subsequent activities, composition of travel party, etc.) were not included in this model, because this model is less interested in the activity patterns and attributes of a choice process involving mode than in the psychological correlates involved in the place, which have an effect on the mode. A full-blown choice model would improve the accuracy of parameter estimates and the overall contribution of sense-of-place indicators; this study provides a framework and rationale for a larger research effort.

To further explore the use of sense-of-place indicators, an analysis of arrival time at the location was conducted to determine which aspects of sense of place were important for explaining variation in time allocation and activity planning. Results of this model are given in Table 5. The dependent variable (arrival time) is reported in minutes.

Results of this regression model indicate that several sense-of-place indicators are significant in explaining variation among respondents. As indicated in the regression model, the intercept of the model is approximately 10:30 a.m. Being male, being employed full time, and having one, two, or three or more cars in the household all contribute to a later arrival to the location. With the exception of being male, all these indicators contribute approximately 1 h to the arrival time. This could reflect that those who are on their lunch break come to the location for food or to run errands. This theory can be further supported by the significance (and contribution of approximately 35 min to the arrival time) of the dummy indicator for a person coming from work. People coming from home are more likely to arrive at the location earlier in the day. Sense-of-place indicators also show significance in explaining arrival time. For instance, variables indicating satisfaction with the entertainment options, no perceived risk of any negative encounters, and self-consciousness elicited while visiting the location all contribute to a later arrival time. Several indicators are significant at a 0.1 level or better. Negative attitudes of parking satisfaction contribute negatively to the arrival time, meaning that people are more likely to come in the morning. People who think these locations lack specific amenities also arrive later in the day.

CONCLUSION

This research was designed to examine the use of sense-of-place attitudinal indicators to explain several aspects of travel behavior. A survey was conducted to collect data about sense-of-place attitudes and travel behavior. Sense-of-place questions were used as indicators in several models of behavior. First, analysis on the two locations and respondents at each location indicate clear differences in socio-demographics as well as sense-of-place attitudes at each location. Models were built to analyze the influence of sense-of-place indicators on the mode used to arrive at each location. Sense-of-place indicators were again found significant and added to the explanatory power of the models. Further testing of the log likelihoods indicating goodness of fit show that the attitudinal place information is of value. The omission of this data in modeling efforts might be detrimental to explaining variation in observed behavior. An additional model examining timing of the activity was conducted for further understanding of how sense-of-place indicators explain observed behavior. Results indicate that sense-of-place indicators again increase the ability of models to capture heterogeneity in observed behavior and can thus predict with higher accuracy.

The research conducted in this study shows promise for the use of sense of place in travel behavior and urban design. Its proven significance and valuable contribution in modeling provides support for further research into its application. Examination of the impact of sense of place in the choice process is a natural direction for research. The interplay of exposure, familiarity, and overall cognition of a place with the development of sense of place should be explored, as this could provide insight into the development of preferences and choices of short-term destination and long-term location choices.

TABLE 5 Regression Model of Arrival Time

Model Variable	Coeff.	SE(HC)	t	$P > \|t\|$
Constant	656.234	24.694	26.575	.000
MALE (1 if male; 0 otherwise)	23.229	11.340	2.048	.041
SBRES (1 if Santa Barbara resident; 0 otherwise)	54.648	14.671	3.725	.000
INCHIGH (1 if annual household income is higher than $100,000; 0 otherwise)				
EMPFULL (1 if employed full-time; 0 otherwise)	48.465	13.561	3.574	.000
EMPPART (1 if employed part-time; 0 otherwise)				
CARS1 (1 if one car in household; 0 otherwise)	51.168	18.551	2.758	.006
CARS2 (1 if two cars in household; 0 otherwise)	49.907	17.727	2.815	.005
CARS3UP (1 if three or more cars in household; 0 otherwise)	46.840	17.293	2.709	.007
WORKDY (1 if work day; 0 otherwise)				
VACDY (1 if vacation day; 0 otherwise)				
HOME_BFR (1 if respondent came from home; 0 otherwise)	−27.466	12.245	−2.243	.025
WORK_BFR (1 if respondent came from work; 0 otherwise)	35.191	17.954	1.960	.050
"I am satisfied with the products offered" (strongly dissatisfied, dissatisfied, slightly dissatisfied, neutral)				
"I am satisfied with the parking" (strongly dissatisfied, dissatisfied, slightly dissatisfied, no opinion)	−27.255	15.543	−1.754	.080
"I am satisfied with the parking" (strongly agree)				
"I am satisfied with the entertainment options" (strongly dissatisfied, dissatisfied, slightly dissatisfied)	32.037	19.470	1.645	.100
"I am satisfied with the entertainment options" (no opinion)	41.478	19.018	2.181	.030
"I am satisfied with the entertainment options" (slightly satisfied)				
"[Loc] makes me feel relaxed" (agree)				
"[Loc] does not involve the risk of unpleasant encounters" (no opinion, slightly agree)	27.932	11.804	2.366	.018
"[Loc] does not make me feel too self-conscious" (strongly dissatisfied, dissatisfied, slightly dissatisfied)	38.174	20.167	1.893	.059
"[Loc] does not lack specific things" (strongly dissatisfied)	40.342	23.912	1.687	.092
"[Loc] does not lack specific things" (dissatisfied, slightly dissatisfied)	25.099	14.187	1.769	.077
"[Loc] does not lack specific things" (no opinion)				

NOTE: $R^2 = .1538$, df = 24, loc = location.

Future research is needed to determine the interplay of sociodemographic cohort and place attitudes and how people choose activity locations.

Integrated transportation land use models have used measurements of accessibility and several physical attributes of a place to explain behavior. These predictors of behavior are a partial measure of the attraction of a place, mode, or trip. This practice can be enhanced with the addition of psychological attributes. The quantification of sense-of-place attitudes, with its rich theoretical support, can provide understanding of the psychological differences in people who gravitate toward different aspects of urban design. The use of sense-of-place factors can introduce additional aspects of design and the latent appeal that these aspects have for different people. Sense-of-place factors also have the potential to further explain reasons for possible successes and failures of land use policies in changing travel behavior. These latent psychological reasons (e.g., positive or negative dispositions) have the potential to explain in greater detail the motivation or lack thereof for various behavioral phenomena. As previously mentioned, the process of changing one's sense of place must also be considered in land use and design policy changes. Expectations and past experiences as well as design elements all weigh in on a person's sense of place, which as seen in this research project have an influence on behavior.

The practicality of this novel variable type must be considered in its application. The prediction of psychological variables is a largely unexplored research area, and presents several difficulties. Future directions of this research include the exploration of significant sociodemographic indicators in predicting latent classes of sense-of-place attitudes. In addition to the sociodemographic indicators, covariates of these indicators could include historical contexts of the individual (such as past exposure to the place, length of residency, research and information regarding the place, etc), or even place information (such as physical characteristics of the place, length of establishment, geographic location, etc), that might also contribute to sense of place. These factors could be used to predict sense-of-place attitudes for individuals similarly to the process of predicting activity generation and scheduling given significant individual attributes. Once these indicators have been predicted, sense of place can be used in activity models to further inform the short term destination choices of individuals as well as medium term choices such as residential or work location choice. Additionally, more general measures of sense of place can be included at a larger aggregation to possibly include as choice attributes. For example, if a general consensus is that a place is associated with being dangerous, this attribute can be included as an indicator to reflect the nature of that detractor for a large percentage of the population in the model.

The interplay of the development of sense of place with experience, the influence of sociodemographic characteristics with sense-of-place development, and the use of sense-of-place attitudes in informing behavior is a complex web of relationships. However, the information provided by these personal details is valuable and has proved effective in enriching the explanatory power of behavior models. Successful quantification and application of this research into several key transportation areas could offer exciting and useful results.

ACKNOWLEDGMENTS

Funding for this research was provided by the University of California Transportation Center and University of California Office of the President. The authors thank the anonymous reviewers for constructive comments and insightful recommendations.

REFERENCES

1. Ettema, D., and H. J. P. Timmermans. *Activity-Based Approaches to Travel Analysis.* Pergamon Press, Oxford, United Kingdom, 1997.
2. Timmermans, H. The Saga of Integrated Land Use-Transport Modeling: How Many Dreams Before We Wake Up? Presented at 10th International Conference on Travel Behavior Research, Lucerne, Switzerland, 2003.
3. Ben-Akiva, M., D. McFadden, T. Gärling, D. Gopinath, J. Walker, D. Bolduc, A. Börsch-Supan, P. Delquié, O. Larichev, T. Morikawa, A. Polydoropoulou, and V. Rao. Extended Framework for Modeling Choice Behavior. *Marketing Letters,* Vol. 10, No. 3, 1999, pp. 187–203.
4. Ben-Akiva, M., D. McFadden, K. Train, J. Walker, C. Bhat, M. Bierlaire, D. Bolduc, A. Boersch-Supan, D. Brownstone, D. S. Bunch, A. Daly, A. De Palma, D. Gopinath, A. Karlstrom, and M. A. Munizaga. Hybrid Choice Models: Progress and Challenges. *Marketing Letters,* Vol. 13, No. 3, 2002, pp. 163–175.
5. Prevedouros, P. D. Associations of Personality Characteristics with Transport Behavior and Residence Location Decisions. *Transportation Research Part A,* Vol. 26, No. 5, 1992, pp. 381–391.
6. Steg, L., C. Vlek, and G. Slotegraaf. Instrumental-Reasoned and Symbolic-Affective Motives for Using a Motor Car. *Transportation Research Part F,* Vol. 4, No. 3, 2001, pp. 151–169.
7. Kostyniuk, L., and D. Eby. Toward a Model of Risky Driving Using Objective Measures of Driving Behavior. Presented at 11th International Conference on Travel Behavior Research, Kyoto, Japan, 2006.
8. Goulias, K., and K. Henson. On Altruists and Egoists in Activity Participation and Travel: Who Are They and Do They Live Together? *Transportation,* Vol. 33, No. 5, 2006, pp. 447–462.
9. Molin, E., T. Arentze, and H. Timmermans. Eliciting Social Network Data. Presented at International Conference on Survey Methods in Transport, Annecy, France, 2008.
10. Clark, A. F., and S. T. Doherty. Examining the Nature and Extent of the Activity-Travel Preplanning Decision Process. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 83–92.
11. Auld, J., A. Mohammadian, and S. T. Doherty. Analysis of Activity Conflict Resolution Strategies. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 10–19.
12. Timmermans, H. Analysis and Models of Household Decision Making Process. Presented at 11th International Conference on Travel Behavior Research, Kyoto, Japan, 2006.
13. Ruiz, T., and M. J. Roorda. Analysis of Planning Decisions During the Activity-Scheduling Process. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 46–55.
14. Auld, J., C. Williams, A. Mohammadian, and P. Nelson. An Automated GPS-Based Prompted Recall Survey with Learning Algorithms. *Transportation Letters,* Vol. 1, No. 1, 2009, pp. 59–79.
15. Kitamura, R., P. L. Mokhtarian, and L. Laidet. A Micro-Analysis of Land Use and Travel in Five Neighborhoods in the San Francisco Bay Area. *Transportation,* Vol. 24, No. 2, 1997, pp. 125–158.
16. Choo, S., and P. Mokhtarian. The Relationship of Vehicle Type, Choice to Personality, Lifestyle, Attitudinal, and Demographic Variables. UCD-ITS-RR-02-06. Institute of Transportation Studies, University of California, Davis, 2002.
17. Kuppam, A. R., R. M. Pendyala, and S. Rahman. Analysis of the Role of Traveler Attitudes and Perceptions in Explaining Mode-Choice Behavior. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1676,* TRB, National Research Council, Washington, D.C., 1999, pp. 68–76.
18. Sunkanapalli, S., R. M. Pendyala, and A. R. Kuppam. Dynamic Analysis of Traveler Attitudes and Perceptions Using Panel Data. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1718,* TRB, National Research Council, Washington, D.C., 2000, pp. 52–60.
19. Dill, J. Travel Behavior and Attitudes: New Urbanist vs. Traditional Suburban Neighborhoods. Presented at 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.
20. Tuan, Y. *Topophilia: A Study of Environmental Perception, Attitudes, and Values.* Prentice-Hall, Englewood Cliffs, N.J., 1974.
21. Golledge, R. G., and R. J. Stimson. *Spatial Behavior: A Geographic Perspective.* Guilford Press, New York, 1997.
22. Bolton, R. *An Economic Interpretation of a "Sense of Place."* Department of Economics, Williams College, Williamstown, Mass., 1989.
23. Canter, D. (ed.). Understanding, Assessing, and Acting in Places: Is an Integrative Framework Possible. Oxford University Press, London, 1991.
24. Proshansky, H. M. The City and Self-Identity. *Environment and Behavior,* Vol. 10, No. 2, 1978, pp. 147–169.
25. Altman, I., and S. M. Low. *Place Attachment.* Plenum Press, New York, 1992.
26. Stokols, D., and S. Shumaker. People in Places: A Transactional View of Settings. In *Cognition, Social Behavior and the Environment* (J. Harvey, ed.), Lawrence Erlbaum, Hillsdale, N.J., 1981.
27. Jorgensen, B., and R. Stedman. Sense of Place as an Attitude: Lakeshore Owners Attitudes Toward Their Propertied. *Journal of Environmental Psychology,* Vol. 21, No. 3, 2001, pp. 233–248.
28. Semken, S., and C. Freeman. Sense of Place in the Practice and Assessment of Place-Based Science Teaching. *Science Education,* Vol. 92, No. 6, 2008, pp. 1042–1057.
29. Lewicka, M. Place Attachment, Place Identity and Place Memory: Restoring the Forgotten City Past. *Journal of Environmental Psychology,* Vol. 28, 2008, pp. 209–231.
30. Smaldone, D., C. Harris, and N. Sanyal. An Exploration of Place as a Process: The Case of Jackson Hole, WY. *Journal of Environmental Psychology,* Vol. 25, 2005, pp. 397–414.
31. Giuliani, M., and R. Feldman. Place Attachment in a Developmental and Cultural Context. *Journal of Environmental Psychology,* Vol. 13, 1993, pp. 267–274.
32. Manzo, L. C. Beyond House and Haven: Toward a Revisioning of Emotional Relationships with Places. *Journal of Environmental Psychology,* Vol. 23, No. 1, 2003, pp. 47–61.
33. Agarwal, P. Operationalizing "Sense of Place" as a Cognitive Operator for Semantics in Place-Based Ontologies. In *Lecture Notes in Computer Science 3693* (A. G. Cohn and D. M. Mark, eds.), Springer, New York, 2005, pp. 96–114.
34. Gustafson, P. Meanings of Place: Everyday Experience and Theoretical Conceptualizations. *Journal of Environmental Psychology,* Vol. 21, No. 1, 2001, pp. 5–16.
35. Hay, R. Sense of Place in Developmental Context. *Journal of Environmental Psychology,* Vol. 18, No. 1, 1998, pp. 5–29.
36. Deutsch, K. *What's So Special About This Place? An Examination of Sense of Place and Its Correlation to Travel Behavior.* MA thesis. University of California at Santa Barbara, 2008.
37. Brown, B. B., and C. M. Werner. Before and After a New Light Rail Stop: Resident Attitudes, Travel Behavior, and Obesity. *Journal of the American Planning Association,* Vol. 75, No. 1, 2009, pp. 5–12.
38. Deutsch, K., and K. Goulias. Investigating the Impact of Sense of Place on Travel Behavior Using an Intercept Survey Methodology. Presented at 88th Annual Meeting of the Transportation Research Board, Washington, D.C., 2009.
39. Stedman, R. Is It Really Just a Social Construction? The Contribution of the Physical Environment to Sense of Place. *Society and Natural Resources,* Vol. 16, No. 8, 2003, pp. 671–685.

The study described in this paper is not a policy or regulation of any public or private agency or organization. Any errors are the sole responsibility of the authors.

The Traveler Behavior and Values Committee peer-reviewed this paper.

Vehicle Emission Control Strategies and Public Opinion in Bhutan

Sherub Phuntsho and Kunnawee Kanitpong

Because of the increasing number of vehicles in Bhutan, vehicle emissions have become a major cause of air pollution in the city of Thimphu. Although the Bhutanese government has implemented many strategies to control vehicle emissions, the strategies appear inadequate; there is still a need to explore alternative strategies for alleviating the problem. Although vehicle emission control strategies can be easily proposed, they are difficult to implement because they depend on public acceptance and support. A questionnaire survey of public opinion on vehicle emission control strategies was carried out in Thimphu City. Data from the survey have been used to analyze individual preference for vehicle emission control strategies and whether they can be statistically linked to socioeconomic characteristics of survey respondents. The results from this study could be helpful to policy makers for evaluating long-term plans to reduce vehicle emissions in Thimphu City.

Poor air quality due to vehicle emissions is a major environmental problem in many cities worldwide. Rapid population growth and increasing motorization generate traffic congestion, resulting in more fuel consumption and, subsequently, air pollution, especially in urban areas. Bhutan, a small, landlocked country located in the center of the eastern Himalayas, whose developmental activities are guided by the concept of gross national happiness (GNH), is experiencing rapid economic growth with an average national growth rate of 9.6% exceeding the target growth rate of 8.2% for the ninth plan period from 2002 to 2008 (*1*). One of the four pillars of GNH is the conservation of the environment and sustainable use of natural resources (*2*).

In Bhutan, rapid economic growth has led to a significant increase in the vehicle fleet in recent years. Vehicle registration increased by more than 17% from 1997 to 2004 (*3*). Recent statistics on vehicle population show that 58% of the total vehicle fleet is in Thimphu City, the capital of Bhutan. Emissions from the growing vehicle population also are increasing, and the trend is likely to continue.

In 1999, the first comprehensive survey on air pollution in Thimphu City was conducted by the Thimphu National Environment Commission (NEC) (*4*). Although the report indicates that vehicle emissions were not a major problem at that time, emissions are becoming a main concern as the number of vehicles rapidly increases. During 1995 and 1996, the emission levels of more than 1,060 vehicles were inspected to establish a vehicle emission standard for Bhutan. It was observed that 66% of petrol vehicles and 96% of diesel vehicles did not meet the Indian emission standard. According to the NEC report, the increase in air pollution due to vehicle emissions is caused by poor quality of fuel, inefficient fuel combustion, and increased traffic movement in Thimphu City (*5*). The study recommended that the government initiate strategies such as establishment of emission standards for vehicles to improve air quality in the city (*3*).

In other countries, many studies have been conducted to propose strategies to reduce and control vehicle emissions (*6–11*). In Bhutan, because of a lack of technical know-how and human resources, extensive research has not yet been carried out. However, the Bhutanese government has implemented many strategies to control vehicle emissions, such as monitoring of fuel quality by banning consumption of leaded gasoline, banning of imported used cars, and establishment of vehicle emission standards in the country. Although many strategies have been proved to be successful, the existing controlled strategies appear to be inadequate. Therefore, there is a need to study other effective strategies to alleviate the problem. Although vehicle emission control strategies can be easily proposed, controls are difficult to successfully implement because they depend on public acceptance and support. Therefore, a study of public opinion is necessary before any proposed strategy is implemented. Public consultation could raise people's confidence and help them believe that there will be no infringement on the rights of individuals in their democratic society. From a policy perspective, it is essential to formulate appropriate policies that could cater to the needs of the target group.

This paper examines public opinion and the acceptability of existing and possible vehicle emission control strategies. A questionnaire survey of public opinion on vehicle emission control strategies was carried out in the city of Thimphu. A statistical technique was applied to the survey data to establish a link between the socioeconomic characteristics of survey respondents and individual preferences for vehicle emission control strategies, which could help to identify particular segments of the population likely to be supportive or against implementation of specific measures.

DATA COLLECTION AND QUESTIONNAIRE SURVEY

Data collection was conducted by questionnaire survey in Thimphu City, the most populated and fastest growing of Bhutan's cities. Questionnaires were randomly distributed to employees of various organizations and to individuals living in Thimphu. The respondents were asked to fill out the questionnaire forms, and the forms were collected after 2 or 3 days. The respondents were asked to answer all questions honestly, including questions related to socioeconomic characteristics and attitudes toward the listed vehicle emission control strategies.

The questionnaire was designed in a simple format to be easy for respondents to understand. The questionnaire was divided into three parts. Questions about socioeconomic characteristics and driving characteristics were included in the first and second parts. In the

Transportation Engineering Program, Asian Institute of Technology, P.O. Box 4, Klong Luang, Bangkok, Thailand 12120. Corresponding author: K. Kanitpong, kanitpon@ait.ac.th.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 103–110.
DOI: 10.3141/2157-13

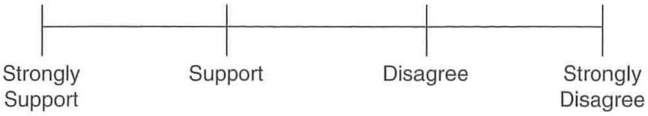


FIGURE 1 Four-point Likert scale.

third part, a Likert scale was used to obtain preference ratings that can quantitatively estimate public opinion. The rates obtained from the Likert scale were then analyzed by assigning a fixed weight on each characteristic of response and then aggregating a total score for a specific group of respondents. The respondents were asked to rate 15 vehicle emission control strategies by using the four-point Likert scale, as shown in Figure 1.

In this part, the questions were given in two categories, as shown in Table 1 (*12–25*). Category A included existing control strategies that have been implemented in Bhutan but require improvement to achieve the goal of vehicle emission reduction. Category B focused on strategies that can be introduced for future implementation. Category B had three subgroups: fiscal measure, infrastructure development, and mandatory regulations. The strategies proposed in Category B were based on an extensive literature review, and some had been proved as successfully implemented strategies in other countries. The strategies are summarized in Table 1.

The survey was conducted during the last quarter of 2008. The questionnaire was randomly distributed to government offices, corporate and private enterprises, schools, and training institutions and in public locations. A total of 790 people in Thimphu were asked

TABLE 1 Vehicle Emission Control Strategies Rated by Survey Respondents

Question No.	Strategies Description
A. Existing Strategies	
A1	Enforce annual emission testing strictly.
A2	Increase the number of city bus services. At least two buses per hour on each existing route operated under city bus services are required (current number of existing bus services varies from 30–90 minutes on nine routes). If the bus service system is improved by reducing the waiting time between the buses, road users will shift their mode choices and the number of private-car users will be reduced.
A3	Convert from two-way narrow streets to one-way streets in the city area. As the two-way narrow streets could create heavy traffic congestions, especially during peak periods, thus generating more exhaust emission from vehicle engines, changing to one-way streets can improve traffic flow and reduce traffic congestion. Emissions could then be reduced because of better traffic flow movement.
A4	Permit the use of some specific models and type of vehicles to be imported to the country. As there are no vehicle manufacturers in Bhutan, all vehicles imported into the country need to meet certain standards of emissions control (*12*). This strategy has been demonstrated as a successful approach to reduce vehicle emissions (*6*).
B. Proposed Strategies: Fiscal Measure	
B1	Increase sales tax and import tax by more than 50% on used passenger cars. By increasing the sales tax and import tax, customers have to pay more costs to own cars. The number of imported used passenger cars that normally generate more pollution is expected to decrease, and hence emissions could be reduced (*6*).
B2	Increase interest rates of vehicle loans for nonpublic-transport vehicles. The recent policy in Bhutan of granting low-interest rates on vehicle loans from financial institutions has increased the purchase rate of private cars in Bhutan. By increasing the interest rates, the number of private cars could be controlled.
B3	Increase vehicle ownership annual expenses (for example: registration renewal fee, road worthiness certificate fee, emission testing fee, etc.). If government increases vehicle ownership taxes, vehicle owners will hesitate to own cars and perhaps be obliged to use the public transport, which can indirectly control emissions. Many European and Asian countries implemented this policy to reduce emissions and congestion (*13, 14*). This strategy has been successfully implemented in Singapore (*15*). Faiz and Dulal (*16*) report that by using this strategy, motor vehicle growth rate was reduced from 4.2% in 1990 to 2.8% in 2002 in Singapore.
B. Proposed Strategies: Infrastructure Development	
B4	Provide electric tram service (green transport) within the city area. In Bhutan, electricity can be produced in tremendous amounts. This is therefore one of the alternatives to provide very clean mass rapid transit in Thimpu City.
B5	Provide bicycle lanes in the city area and promote the use of bicycles for short trips. This strategy will limit the use of private cars in the central business district (CBD) (*17*). The study by Petritsch et al. (*18*) shows that the addition of cycling facilities can result in 5% to 10% mode shift from motorized to nonmotorized transport.
B6	Improve pedestrian walkways in the CBD. Improving walkability is another strategy to promote and encourage nonmotorized transportation use, which can be effective in emission reductions by reducing short motorized trips (*17*). ADONIS, Litman et al., and Litman suggest many ways to improve walkability, such as improving walkways and other walking facilities (*19, 20, 21*).
B7	Construct pedestrian footbridges at busy street crossings within the city area. This strategy uses a concept similar to B6, which will improve the walking facilities for pedestrians in the city area.
B. Proposed Strategies: Mandatory Regulations	
B8	Retrofit and scrap motor vehicles that have emissions beyond the standards. When vehicles are older and release emission levels that exceed the acceptable limits, they should be inspected within a period of time. If the vehicles still fail the legal requirements within the given time, they must be retrofitted or scrapped. This policy has been implemented in the United States, Australia, and some European Union countries (*6, 7, 9, 17, 22, 23*).
B9	Switch bus engines from diesel fuel to compressed natural gas (CNG) for public transport vehicles. Alternate fuels such as CNG can be used to replace diesel because of its clean burning characteristic and very low amount of exhaust pollution. Many countries have successfully implemented this strategy (*6, 7, 9, 24*).
B10	Establish a compulsory campaign of No-Private-Vehicle Day (only walking to office or work). The No-Private-Vehicle Day campaign has been implemented voluntarily in Bhutan since 2008. In this study, it is suggested as a compulsory campaign. The bus service should be operated to provide services for long trips.
B11	Install on-board diagnostic (OBD) systems in vehicles to alert drivers when the emissions from their vehicles exceed the limits. OBD systems are designed to monitor the performance of an engine's major components, including those responsible for controlling emissions (*7, 25*).

TABLE 2 Socioeconomic Characteristics of Survey Respondents

Category	Description	Percent
Gender	Male	67.3
	Female	32.7
Age (years)	<20	8.4
	20–30	69.3
	31–40	19.0
	>40	3.3
Occupation	Government employee	31.5
	Private employee	1.8
	Corporate employee	38.6
	Business	1.3
	Police or soldier	2.0
	Housewife	8.0
	Trainee or student	24.1
Education	Primary or lower	8.6
	High school	43.9
	Diploma	26.6
	Bachelor's degree	16.0
	Master's degree	4.6
	Doctoral degree	0.3
Monthly income (in ngultrum) (1 US$ = 47.29 ngultrum)	<5,000	34.0
	5,000–10,000	28.4
	10,001–15,000	25.4
	15,001–20,000	7.9
	20,001–30,000	3.8
	>30,000	0.5
Car ownership	Yes	33.0
	No	67.0
Mode of transportation	Private car	37.3
	Taxi	30.8
	City bus	31.0
	Two-wheeler	3.3
Trip purpose	Office or other work	58.1
	Business	3.6
	Education (self or children)	14.2
	Shopping	16.8
	Recreation	5.6
	Other	1.8
Duration of stay (years)	<1	8.9
	1–2	13.7
	2–3	16.5
	>3	60.9
Driving license	Yes	37.31
	No	62.69

to complete the questionnaire. Only 394 samples were used in the analysis, because about half the respondents appeared to have very little knowledge about vehicle emission problems in Bhutan, and a number of incomplete questionnaires were excluded. Table 2 describes the socioeconomic characteristics of the respondents.

PUBLIC OPINION OF STRATEGIES

Descriptive Analysis

A total of 15 vehicle emission control strategies were rated by respondent opinion. These strategies were in two categories: existing strategies and proposed strategies. The existing strategies are those already implemented but needing improvement or strict enforcement of regulations. The proposed strategies include various measures that could be feasible to implement and cost-effective in the Bhutanese context. The proposed strategies are subdivided into three groups: fiscal measure, which has financial implications for people who contribute to vehicle emission pollution; infrastructure development, whereby the government or concerned agencies are required to bring development through infrastructure buildings; and mandatory regulations, in which the government or concerned agencies are required to put in place a strong legal framework for uncompromised implementation of regulations.

Results from descriptive analysis of public opinion about existing vehicle emission control strategies are presented in Figure 2. It was found that 88% to 94% of the respondents agreed with the strict enforcement of mandatory annual emission testing and an increase in the number of city bus services. Fewer people gave their support to the conversion from two-way narrow street to one-way street in the city. Short interviews with some respondents suggested that this strategy unnecessarily increases travel distance and, more important, causes confusion to road users. Also, respondents were aware that longer travel distance could increase pollution in the city. Yet, this strategy was found to be supported by the majority of the respondents. Almost 50% of the respondents were against a policy permitting the use of some specific model and type of vehicles to be imported into the country. They argued that the specific model and type of imported vehicles are expensive and could affect the affordability of cars for low-income people. The respondents also expressed some concerns regarding the social and economic gaps between rich and poor people.

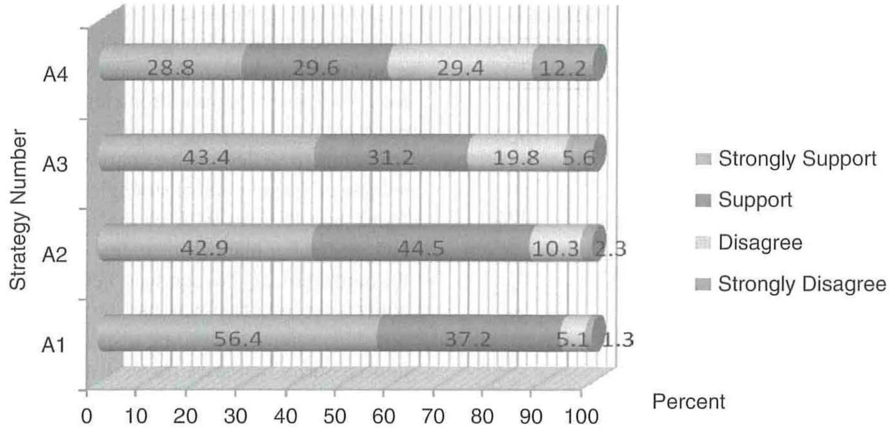
Figure 3 shows the results of public opinion on the proposed vehicle emission control strategies. Generally, the proposed strategies are well supported by the public, except for fiscal measures B1, B2, and B3, which are the least favorable option in public opinion. This finding is as expected because those strategies financially affect those who plan to buy car and those who already own car. Among the proposed strategies in the infrastructure development group, B5—the strategy to provide bicycle lanes in the city area and to promote the use of bicycle for short trips—has been given relatively less support by respondents. The improvement of pedestrian walkways (B6) and the construction of pedestrian footbridges (B7) were strongly supported by the public. It is evident that most respondents prefer improvement of pedestrian facilities. For mandatory regulations, it was found that B8, B9, B10, and B11 were all strongly supported by respondents.

Preferential Ranking of Strategies

The preference responses of the respondents were analyzed to evaluate public opinion on the existing and proposed vehicle emission control strategies in the quantitative measures. The rates obtained from the Likert scale were analyzed by assigning a fixed weight on each response and summing individual scores to determine the total score. The total scores were used to represent the level of preference responses of the respondents. In this study, it is assumed that the weights associated with the responses are equivalent to the values of 2, 1, −1, and −2, which are assigned to strongly support, support, disagree, and strongly disagree, respectively. Thus, the higher the total score, the higher the preference the public gave to the emission control strategy. Response ratings are summarized in Table 3.

Given vehicle emission control strategies currently implemented in Bhutan, the results obtained from the survey indicate that strict enforcement of annual emission testing is the most desirable strategy for controlling vehicle emissions in the city. The increase in the number of city bus services is found to be the second most popular strategy because it can help to reduce the number of private cars used in the city area and hence lower the vehicle emissions rate.

For the proposed strategies, the survey indicates that the group of mandatory regulations is the most popular strategy and hence a viable option for the future, followed by those involving infrastructure

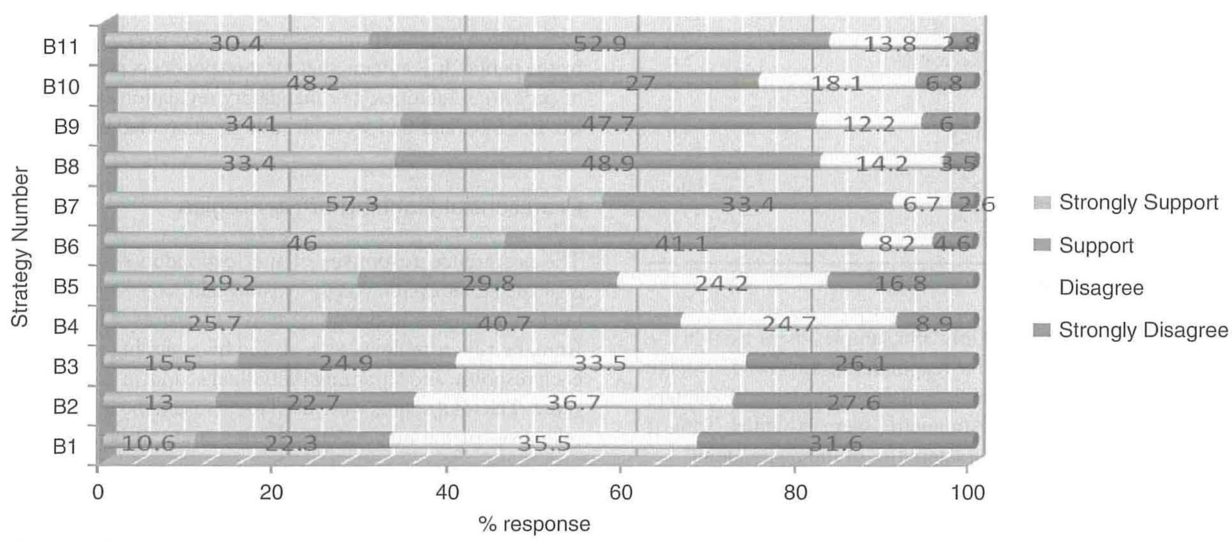


A1: Strict enforcement of annual emission testing.
A2: Increase the number of city bus service.
A3: Convert from two-way narrow street to one-way street in the city area.
A4: Permit the use of some specific model and type of vehicles to be imported to the country.

FIGURE 2 Public opinion on existing vehicle emission control strategies.

development and fiscal measures. Among the mandatory regulations is the strategic idea of retrofitting and scrapping motor vehicles with a level of emissions exceeding the standard that is rated with the highest score on the Likert scale, although the scores of all other regulations are slightly lower.

The provision of pedestrian footbridges at busy street crossings is ranked as the most popular strategy involving infrastructure development, followed by improving pedestrian walkways in the central business district (CBD). These findings could reflect the need for better pedestrian facilities in the sense that policies to promote a better walking environment could lead to changes in mode of transportation for short trips and minimize the number of motorized trips and hence reduce vehicle emissions in the city. Provision of electric tram and bicycle lanes in the city area are the two least-popular



B1: Increase sales tax and import tax by more than 50% on used passenger cars.
B2: Increase interest rates of vehicle loan for non-public transport vehicles.
B3: Increase vehicle ownership expenses annually
B4: Provide electric tram (green transport) within the city area.
B5: Provide bicycle lanes in the city area and promote the use of bicycle for short trip.
B6: Improve pedestrian walkways in CBD area.
B7: Construction of pedestrian footbridges at busy street crossings within the city area.
B8: Retrofit and scrap motor vehicles which have emissions beyond the standards.
B9: Switch from diesel fuel to Compressed Natural Gas (CNG) bus engine for the public transport vehicles.
B10: Compulsory campaign of "No Private Vehicle Day"
B11: Install On-Board Diagnostic (OBD) system in the vehicles to alert drivers when the emission from their vehicle exceeds the limits.

FIGURE 3 Public opinion on proposed vehicle emission control strategies.

TABLE 3 Preferential Rankings of Vehicle Emission Control Strategies

	Existing Strategies				Proposed Strategies										
					Fiscal Measure			Infrastructure Development				Mandatory Regulations			
Strategy	A1	A2	A3	A4	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11
Average score	1.426	1.149	0.859	0.358	−0.529	−0.395	−0.121	0.528	0.269	1.173	1.319	1.003	0.911	0.868	0.956
Measure average						−0.348			0.822					0.935	

strategies in this group. However, the majority of people express a positive attitude toward both strategies, as indicated by the positive scores on the Likert scale. All fiscal-measure strategies including an increase in sales tax and an import tax for used passenger cars, an increase in interest rates for vehicle loans for non-public-transport vehicles, and an increase in the annual fee for vehicle ownership are rated poorly on the Likert scale, having negative total scores. This is not surprising, as most fiscal policies adding to people's expenses tend to receive low public support.

FACTORS AFFECTING INDIVIDUAL PREFERENCES FOR STRATEGIES

Ordered Logit Regression Model

To evaluate the significant factors affecting the preference rate of vehicle emission control strategies, the ordered logistic regression technique was applied in this study. Given the public opinion on the Likert scale, an observed rating for emission control strategies is an indicator of the utility distribution. The data obtained from the survey were analyzed by using ordered logit models to determine the factors that influence the choice process of individuals in the context of emission control strategies. The dependent variable in this regression model is the response rated from the Likert scale, with four orders defined as −2 for strongly disagree, −1 for disagree, 1 for support, and 2 for strongly support. The independent variables considered in the analysis, as summarized in Table 4, include individual characteristics of the respondents, such as gender, age, occupation, education, income, car ownership, mode of transportation, trip purpose, duration of stay in Thimphu City, and driving license. However, the test of multicollinearity indicates the existence of strong correlation among three variables: transport mode, car ownership, and driving license (i.e., all pairwise correlation coefficients are higher than 0.6). The latter two variables are therefore excluded in the preferred model specification. Table 5 presents estimation results from the ordered logit models. The relative magnitude of estimated coefficients indicates the extent to which socioeconomic and driving characteristics affect individual preferences to vehicle emission control strategies in Thimphu City.

Results for Existing Strategies

A1. Strict Enforcement of Annual Emissions Testing

It was found that the coefficients of the independent variable education is statistically significant at the 5% level, and those of the variables income and duration of stay are statistically significant at the 10% level. These findings suggest that highly educated respondents strongly support the strategy of strict enforcement of annual emission testing of all vehicles, whereas higher-income respondents and respondents who stay in Thimphu City for longer than 3 years also support this strategy.

A2. Increase City Bus Service

The respondents whose trip purpose was shopping or recreational activities significantly supported increasing city bus frequency at the 5% level.

A3. Convert from Two-Way Narrow Streets to One-Way Streets in City Area

None of individual characteristics of respondents was found to explain attitudes toward the conversion from two-way narrow streets to one-way streets in the city area.

TABLE 4 Independent Variables

Variable	Definition
GENDER	Gender (1 if the respondent is female, 0 otherwise)
AGE	Age (continuous variable)
GOVERNMENT	Government employee (1 if the respondent is government employee, 0 otherwise)
PRIVATE	Private employee (1 if the respondent is corporate or private employee, 0 otherwise)
STUDENT	Student [1 if the respondent is trainee or student, 0 otherwise]
EDUCATION	Education [1 if the respondent is college educated (diploma and higher education), 0 otherwise (high school and lower education)]
INCOME	Income (1 if the respondent has income >15,000 Nu^2, 0 otherwise)[a]
CAR OWNERSHIP	Car ownership (1 if the respondent owns car, 0 otherwise)
MODE	Mode (1 if the respondent commutes by private car, 0 otherwise)
WORK TRIP	Work or office trip purpose (1 if the trip is for work or office purpose, 0 otherwise)
EDU TRIP	Educational trip purpose (1 if the trip is for educational purpose, 0 otherwise)
SHOPPING TRIP	Shopping or recreational trip purpose (1 if the trip is for shopping or recreational purpose, 0 otherwise)
DURATION	Duration of stay (1 if the respondent stays in the study area more than 3 years, 0 otherwise)
LICENSE	Driving license (1 if the respondent has a driving license, 0 otherwise)

[a]Nu^2 = monthly income in ngultrum.

TABLE 5 Coefficients of Ordered Logit Model for Vehicle Emission Control Strategies

	Existing Strategies				Proposed Strategies		
					Fiscal Measure		
Variable	A1	A2	A3	A4	B1	B2	B3
GENDER	−0.062	−0.253	−0.154	−0.233	0.136	0.253	0.042
AGE	0.002	0.009	0.013	0.007	−0.019	0.017	−0.013
GOVERNMENT	0.102	−0.169	0.237	1.018*	−0.386	−0.775	−0.028
PRIVATE	0.118	−0.195	−0.062	0.892*	−1.060**	−1.580**	0.055
STUDENT	−0.723	−0.058	−0.498	0.992*	0.148	−0.137	0.064
EDUCATION	0.769**	0.144	−0.113	−0.098	0.199	0.651**	−0.231
INCOME	0.853*	−0.397	−0.252	−0.105	0.208	−0.042	−0.356
MODE	−0.005	−0.223	0.151	−0.097	0.323	0.222	−0.706***
WORK TRIP	0.295	0.626	0.305	0.054	−0.148	0.216	−0.014
EDU TRIP	0.556	0.366	0.222	0.241	−0.213	0.148	−0.028
SHOPPING TRIP	0.298	1.037**	−0.109	0.078	−0.553	0.133	0.516
DURATION	0.416*	0.002	−0.086	0.096	−0.117	−0.426**	0.110
Log likelihood	−331.04	−407.89	−465.69	−522.63	−499.45	−505.28	−525.69
No. of observation		394				394	

***Significance at the 1% level, **significance at the 5% level, *significance at the 10% level.

A4. Permit Importation of Specific Model and Type of Vehicles

The coefficients of the occupation variables (government employee, private employee, student) are found to be statistically significant at the 10% level. The highest positive sign of 1.018 for the government employee variable indicates that government officers are more likely to support the use of some specific models and types of vehicles to be imported into the country.

Results for Proposed Strategies: Fiscal Measure

B1. Increase Sales and Import Tax by More Than 50% on Used Passenger Cars

The negative coefficient of private employee is statistically significant at the 5% level, implying that corporate or private employees do not agree with the strategy of increasing sales tax and import tax by more than 50% on used passenger cars.

B2. Increase Interest Rates for Loans for Non-Public-Transport Vehicles

Corporate or private employees and the respondents who stay in Thimphu City for longer than 3 years significantly disagree with the increase of interest rates on loans for non-public-transport vehicles. In contrast, it was found that highly educated people support this policy.

B3. Increase Annual Vehicle Ownership Expenses

The negative and significant coefficient for the mode of travel suggests that most respondents who usually commute by private car are strongly against the strategy of increasing the annual expenses of vehicle ownership. This is what one would expect since the strategy directly affects their living costs.

Results for Proposed Strategies: Infrastructure Development

B4. Provide Electric Tram Service Within City Area

Although the finding that government and private employees strongly oppose the provision of an electric tram service within the city area is somewhat difficult to explain, highly educated respondents and respondents who commute by private car were found to strongly support this strategy. The latter finding is somewhat as expected since the electric tram both is environmentally friendly and provides an alternative means of travel in the city.

B5. Provide Bicycle Lanes in City and Promote Use of Bicycles for Short Trips

Highly educated respondents and those living in Thimphu City for longer than 3 years strongly support the strategy of providing bicycle lanes in the city area and promoting the use of bicycles for short trips. However, it was found that female respondents strongly oppose the idea of promoting bicycle travel. This is not surprising because this strategy is not beneficial to women, whose traditional Bhutanese dress, called *kira* and regularly worn in daily life, is not suitable for bicycle riding.

B6. Improve Pedestrian Walkways in CBD

The coefficients for all occupation variables are found to be statistically significant at the 10% level, implying that government and private employees as well as students give their support to the improvement

Infrastructure Development				Mandatory Regulations			
B4	B5	B6	B7	B8	B9	B10	B11
−0.278	−0.690***	−0.245	−0.169	−0.485**	0.009	−0.074	−0.241
−0.015	0.009	0.019	0.022	0.016	0.019	0.035*	−0.021
−1.316**	0.366	0.898*	−0.422	−0.043	0.201	0.549	1.588**
−1.244**	0.675	1.027*	−0.524	0.131	0.123	0.372	1.571**
−0.780	−0.313	1.099*	−0.806	−0.065	−0.358	0.333	1.077*
0.569**	0.631**	0.199	0.173	0.878**	0.243	−0.033	0.303
−0.440	−0.031	−0.052	0.060	0.298	0.058	−0.760**	−0.008
0.364*	−0.216	0.142	0.063	0.025	0.151	−0.052	0.160
0.146	0.273	−0.318	0.350	1.157**	−1.393**	0.433	0.179
0.304	0.756	−0.706	0.599	0.964*	−1.025*	0.119	0.733
0.199	0.306	0.050	0.854*	0.751	−0.923*	0.447	0.203
0.239	0.369*	−0.083	0.299	−0.030	−0.327	−0.575**	0.263
−489.93	−516.73	−413.61	−377.57	−423.57	−444.10	−466.76	−411.38
	394					394	

of pedestrian walkways in the CBD. The largest magnitude of the student coefficient (1.099) suggests that students are most likely to support this strategy.

B7. Construct Pedestrian Footbridges at Busy Street Crossings Within City

It was found that respondents strongly support the construction of footbridges when traveling for shopping or recreational activities. This could be because most shopping areas and recreational centers are located in the city area. People could easily cross a road and walk around shopping and recreational areas without the need to use any public transport or private car.

Results for Proposed Strategies: Mandatory Regulations

B8. Retrofit and Scrap Motor Vehicles that Have Emissions Beyond Standards

Highly educated respondents and those with the trip purpose of going to work or to school are found to strongly support the strategy. However, the coefficient for the gender variable is negative and statistically significant at the 5% level. The gender difference in attitude toward this strategy is unexpected, and it is difficult to provide a plausible explanation.

B9. Switch Public Transport Vehicles from Diesel Fuel to Compressed Natural Gas

The negative coefficients of all trip purpose variables are statistically significant at conventional levels. However, there appears to be no valid assumption about the relationship between the trip purpose characteristics of respondents and the opinion on the strategy to switch bus engines from diesel fuel to compressed natural gas.

B10. Compulsory Campaign for No-Private-Vehicle Day

As shown in Table 5, the positive coefficients for age are found to be statistically significant at the 5% level, meaning that older respondents strongly support the campaign for a no-private-vehicle day. However, high-income respondents and those having lived in Thimphu City longer than 3 years do not support the campaign. Some respondents who have lived in Thimphu City longer than 3 years mentioned that this compulsory campaign could adversely affect business and economic development of the country. In addition, it is unsurprising to see that higher-income people do not support the campaign because they usually travel by their own car and might be unwilling to change to other transportation modes.

B11. Install On-Board Diagnostic System in Vehicles to Alert Drivers to Excessive Emissions

The coefficients of all occupation variables were found to be statistically significant. Government or private employees and students strongly support the installation of on-board diagnostic systems in vehicles. The highest positive coefficient of 1.588 for the government employee variable implies that government officers are most likely to support this strategy.

SUMMARY AND DISCUSSION

This paper attempted to examine public opinion and the acceptability of existing and some possible vehicle emission control strategies aimed at alleviating the air pollution problem in Thimphu, the capital city

of Bhutan. The methodology used in this study also provides an approach to quantify the relative preferences of different groups of people and their attitudes toward policy decisions, which could facilitate the decision-making process for selecting appropriate strategies for predetermined target groups.

Of the existing implemented strategies in Bhutan, results from the survey suggest that the strict enforcement of annual emission testing is the most favorable strategy for vehicle emissions control. The second most popular strategy is the increase in city bus services, which can reduce the number of private cars used in the city area and hence lower the rate of vehicle emissions. Of the proposed strategies, the group of mandatory regulations was found to be the most popular strategy, followed by policies involving infrastructure development and fiscal policy measures.

Several groups of people in Thimphu appear to give their support to implementation of vehicle emission control strategies. However, there is disagreement by specific groups of people who expressed opinions against some strategies, which must be taken into account in policy formulation and implementation. Ordered logit model estimation reveals that highly educated people are supportive to the existing policy of strict enforcement of annual emission testing. They were found to be favorable to a policy of retrofitting or scrapping vehicles that generate emissions exceeding the standard level, and to controlling the number of private cars by increasing the interest rate for vehicle loans and promoting bicycle use.

People with private cars support the policy of promoting the use of public transportation in the city area. Yet they are against increasing vehicle ownership expenses. This reflects public attitudes favoring transport infrastructure supply over transport demand control through use of fiscal measures. Similar results were found for private employees, who strongly disfavor any fiscal measure on private car ownership. High-income people are not opposed to any fiscal measure strategies but are not supportive of the compulsory campaign for a no-private-vehicle day. Because of the discomfort and inconvenience of trips without private cars, high-income people expressed opinions against this campaign.

The results also reveal that government employees, private employees, students, and people whose trips are for shopping or recreational purposes support improvement of the pedestrian infrastructure. This implies a possibility of mode shift from motorized to nonmotorized transport, which could help somewhat to reduce vehicle emissions if policy makers provide safe and agreeable conditions for pedestrian facilities. To promote bicycle use by providing bicycle lanes, public agencies in Thimphu need to consider the gender issue. This study found evidence that the policy does not appear to be beneficial to women, possibly because of the inconvenience of riding bicycle while wearing traditional dress.

REFERENCES

1. Ninth Plan a Success: Government. *Bhutantimes,* Jan. 2, 2009.
2. Champion of the Earth. *Kuensel,* April 23, 2005.
3. *Bhutan Environment Outlook 2008.* National Environment Commission, Thimphu, Bhutan, 2008.
4. *Environmental Assessment Act 2000.* National Environment Commission, Thimphu, Bhutan, 2000.
5. *The Middle Path: National Environment Strategy of Bhutan.* National Environment Commission, Thimphu, 1998.
6. *Older Gasoline Vehicles in Developing Countries and Economies in Transition: Their Importance and the Policy Options for Addressing Them.* Organisation for Economic Cooperation and Development, Paris, 1999.
7. Faiz, A., K. Sinha, M. Walsh, and A. Varma. *Automotive Air Pollution: Issues and Options for Developing Countries,* World Bank, Washington, D.C., 1990.
8. Faiz, A., C. S. Weaver, and M. Walsh. *Air Pollution from Motor Vehicles: Standards and Technologies for Controlling Emissions.* World Bank, Washington, D.C., 1996.
9. Krupnick, A. J. Transportation and Urban Air Pollution Policies for Developed and Developing Countries. In *Transportation Research Record 1312,* TRB, National Research Council, Washington, D.C., 1991, pp. 90–98.
10. Acutt, M. Z., and J. S. Dodgson. Controlling the Environmental Impacts of Transport: Matching Instruments to Objectives. *Transportation Research Part D,* Vol. 2, No. 1, 1997, pp. 17–33.
11. Timilsina, G. R., and H. B. Dulal. *Fiscal Policy Instruments for Reducing Congestion and Atmospheric Emissions in the Transport Sector: A Review.* Policy Research Working Paper 4652. World Bank, Washington, D.C., 2008.
12. *National Environment Protection Bill of Bhutan 2007.* Royal Government of Bhutan, Thimphu, 2007.
13. Kunert, U., and H. Kuhfeld. The Diverse Structures of Passenger Car Taxation in Europe and the EU Commissions Proposal for Reform. *Transport Policy,* Vol. 14, No. 4, 2007, pp. 306–316.
14. Hirota, K., J. Poot, and K. Minato. Do Policy Incentives Affect the Environmental Impact of Private Car Use? Evidence from a Sample of Large Cities. Presented at 43rd Congress of the European Regional Science Association, Finland, 2003.
15. Chin, A. T. H. Containing Air Pollution and Traffic Congestion: Transport Policy and the Environment in Singapore. *Atmospheric Environment,* Vol. 30, No. 5, 1996, pp. 787–801.
16. Faiz, A., and H. B. Dulal. *Policy Instruments for Reducing Transport Pollutant and GHG Emissions.* World Bank, Washington, D.C., 2009.
17. Energy Conservation and Emission Reduction Strategies. In *TDM Encyclopedia,* Victoria Transport Policy Institute, Victoria, Australia, 2009.
18. Petritsch, T. A., B. W. Landis, P. S. McLeod, H. F. Huang, and D. Scott. Energy Savings from Provision of Bicycle Facilities. Presented at 87th Annual Meeting of the Transportation Research Board, Washington, D.C., 2008.
19. *Best Practice to Promote Cycling and Walking and How to Substitute Short Car Trips by Cycling and Walking.* ADONIS Transport RTD Program, European Union, Brussels, Belgium, 1999.
20. Litman, T., R. Blair, B. Demopoulos, N. Eddy, A. Fritzel, D. Laidlaw, H. Maddox, and K. Forster. *Pedestrian and Bicycle Planning: A Guide to Best Practices.* Victoria Transport Policy Institute, Victoria, Australia, 2006.
21. Litman, T. A. Economic Value of Walkability. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1828,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 3–11.
22. *Transport and Greenhouse: Costs and Options for Reducing Emissions.* Bureau of Transport Economics, Canberra, Australia, 1996.
23. *Accelerated Vehicle Retirement Programs.* Office of Mobile Sources, U.S. Environmental Protection Agency, 1997.
24. Hao, J., J. Hu, and L. Fu. Controlling Vehicular Emissions in Beijing During the Last Decade. *Transportation Research Part A,* Vol. 40, No. 8, 2006, pp. 639–651.
25. *On Board Diagnostics.* U.S. Environmental Protection Agency, 2009.

The Travel Analysis Methods Section peer-reviewed this paper.

Urban Structure and Socioeconomic Barriers to Consumer Adoption of Energy-Efficient Automobile Technology in a Dispersed City

Case Study of Brisbane, Australia

Jago Dodson, Terry Li, and Neil Sipe

The capacity for suburban households to respond to a changing global energy context by changing their motor vehicle technology is examined. Transforming transport systems will make up a crucial element in policy and planning responses to energy and climate challenges. Government policy appears focused on a transition to more-efficient vehicle types or alternative fuel and engine types. Yet such policies have failed to account for the considerable social differences in household exposure to the costs of transport energy and the adaptability of households in altering their use of modes and vehicle types. Nor do such policies recognize how urban social structure, household social status, and automobile type intersect spatially within Australian cities. The links between urban social structure and composition of the motor vehicle fleet are examined to test whether the households that are most reliant on motor vehicles for transport have the financial capacity to rapidly alter their vehicle technology in response to changing energy prices and supply conditions. Australian Bureau of Statistics census data and motor vehicle registration data at the postcode level are used to compare socioeconomic status with the age, fuel consumption, and value of the suburban vehicle fleet for the Brisbane and South East Queensland regions of Australia. This spatial deployment of census and vehicle registration data is novel in the Australian context. It is argued that policies that focus on vehicle technology alone face a number of social equity hurdles as measures to overcome urban transport fuel security problems.

It is clear that the world faces a major change in the type and rate of energy consumption. Between 2004 and 2009, the pattern of world oil prices changed dramatically. Petroleum prices became highly variable, racing from low, reasonably static levels to levels not seen for almost three decades. The price of oil in early 2004 was around US$25 per barrel but by mid 2008 exceeded US$140 per barrel. Higher oil prices translated into much higher fuel costs and caused considerable anxiety among many economically advanced nations with widespread concern about the impact of high oil costs on economic activity. Also, global petroleum production likely will peak as depleting reserves prove incapable of meeting demand.

Accompanying the stresses facing the global petroleum supply system is a heightened awareness of the impact of human carbon emissions on the global climate system. A reduction in climate emissions is established policy among the governments of developed nations. In Australia, the federal government is seeking to institute a trading system for greenhouse gas emissions to reduce national emission levels and is pursuing a range of industry support measures to encourage low-emission production and consumption practices.

The automotive industry faces considerable transformation if it is to respond to the challenges of petroleum depletion and climate mitigation. There is evidence that the industry is poorly positioned to respond to the changing energy environment, as the bankruptcy of the General Motors Corporation attests. Around the world, research into energy-efficient vehicles is accelerating. The Australian government is spending AU$1.3 billion on a green car fund (*1*), which is intended to reduce fuel consumption and greenhouse gas emissions of the passenger vehicle fleet (AU$1.00 = US$0.93).

The pursuit of a green car has largely been an effort to develop and supply a new technology and fuel combination that can deliver current patterns of mobility at comparable levels of affordability. Little discussion has questioned whether implementation of a low-carbon or low-petroleum vehicle fleet is socially achievable. There has been almost no assessment of the differential socioeconomic demand capacity of individuals and households for new vehicle and fuel types. How long will new vehicles take to filter through the secondhand market to become affordable to lower-income groups? Can policies that rely on new vehicle technology meet greenhouse and energy objectives in a timely, socially equitable way?

This paper explores the links between urban social structure and the composition of the urban automobile fleet to assess the social implications of urban transport policies that rely on vehicle and fuel innovation.

ENERGY CHALLENGES

The world is facing major energy challenges (*2*), in particular, managing greenhouse emissions and the depletion of petroleum resources. The sharp increase in oil prices from 2004 to 2008 drew attention to the problems of global petroleum supplies. The oil-price shock was the

J. Dodson and T. Li, Urban Research Program, and N. Sipe, School of Environment, Griffith University, Nathan Campus, Nathan, Queensland, 4111, Australia. Corresponding author: N. Sipe, n.sipe@griffith.edu.au.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 111–118.
DOI: 10.3141/2157-14

result of several factors, including strong global growth and petroleum demand, constraints on production facility capacity, declining production in some major fields, geopolitical uncertainty, speculative investment, and anxiety about future petroleum supply. The rapid growth of China and India, for example, has received considerable attention from both scholars and policy makers for increasing global petroleum consumption (*3, 4*). The growing awareness of global energy challenges has been accompanied by geopolitical uncertainty and indications that major oil-consuming nations are repositioning for an era of greater international competition for petroleum reserves and resources (*5, 6*).

The long-run sustainability of petroleum supplies is also in question. The literature of the last decade suggests that global petroleum production may soon hit a peak of maximum production followed by a decline. The notion of peak oil has been used to describe a scenario in which declining petroleum production is unable to satisfy global demand, leading to increasing prices for petroleum fuels (*7–10*). A peak in global petroleum production by 2020 would likely bring severe petroleum shortages and a consequence of rapidly rising fuel prices. Transport systems that depend on petroleum could face severe and prolonged disruption without widespread deployment of alternative fuel types and vehicle technology or forms of mobility that do not rely on petroleum. In Australia, the transport sector accounts for 36% of national energy consumption from all sources (*11*) and is responsible for more than 70% of national petroleum consumption (*12*). In this context, questions about the capacity of major economies to transform their urban transport systems, including vehicle fleets, to be less reliant on petroleum have major significance for public policy.

The question of petroleum energy security intersects with concerns about the condition and future of the global climate. It is widely accepted by scientific bodies that global climate change likely has been induced by human activity, and considerable policy and political effort is being expended on measures to reduce emissions of greenhouse gases. Transport systems contribute approximately 13.1% of global greenhouse emissions and about 23% of emissions due to energy use (*13*). The Intergovernmental Panel on Climate Change anticipates that light vehicles will produce 38% of the additional 100 exajoules of global transport emissions expected by 2050 (*13*). In Australia, transport emissions are 14% of national emissions, and approximately 53% of all Australian transport emissions are from private passenger vehicles (*14*).

CURRENT AUTOMOTIVE SECTOR POLICY IN AUSTRALIA

More than 97% of Australian transport is powered by petroleum fuels in the form of petrol, diesel, aviation gasoline, or liquefied petroleum gas. Australian cities are among the most petroleum dependent in the world outside North America. Road vehicles accounted for 75% of Australian transport fuel consumption, and passenger vehicles are 62% of this total (*11*). The fleet of private passenger motor vehicles is a key contributor to Australia's petroleum energy dependence and to national greenhouse emissions. Reducing the energy and carbon intensity of the Australian transport system, especially the private motor vehicle fleet, is a key objective for policy makers.

Current Australian government policy for reducing energy dependence and the greenhouse impact of use of private motor vehicles focuses on two strategies. First is an intent to reduce private motor vehicle emissions by improving vehicle fuel efficiency via new drive trains, such as hybrid petrol electric or plug-in electric motors. At the federal level, this policy shift is signified by the Australian government's recently announced $6.2 billion Green Vehicle Plan, which encourages innovation within the domestic automotive sector (*1*).

A second policy objective is to shift to renewable forms of fuel, such as biofuels and electricity. The main energy form being pursued by motor vehicle producers globally is electric vehicles. The Australian government's automotive industry review anticipated an array of electric and hybrid petrol–electric vehicles coming to market in the 2010–2012 period (*15*).

The excitement around proposals for mass conversion of automobile fleets to hybrid or fully electric vehicles may need to be tempered by the historical experience of both improved fuel economy and the provision and uptake of such technology. Australia has a modest record for vehicle fuel-economy improvements with average fleet efficiency remaining effectively unchanged since the early 1960s at around 11.4 L/100 km (*16*). This almost static fuel economy appears to be due to the offset of increasing engine efficiency by increasing vehicle mass and the introduction of energy-consuming accessories such as air conditioning and audio equipment. The change in Australian fuel economy is much weaker than that observed for cars in the United States, which has seen fuel economy increase from approximately 17.6 L/100 km in 1973 to 10.5 L/100 km in 2006 (*17*).

These patterns imply that advances in energy efficiency and improvements in greenhouse gas emissions heralded by electric vehicles may be more complex than proponents admit. But two further problems afflict the proposition that electric vehicles can offer a comprehensive alternative to current vehicle technologies. First, even if fully electric vehicles were introduced by 2012, their anticipated limited range and lack of supporting infrastructure would prevent widespread adoption. Second, petrol–electric vehicles sell at a premium relative to conventional vehicle types and are therefore unattractive to large segments of the vehicle market.

URBAN TRANSPORT PATTERNS AND SOCIOSPATIAL STRUCTURE

The uncertain cost and complex rollout of new vehicle technology will intersect with a set of social and economic processes within Australian cities. These processes will complicate the adoption of new automobile technologies. Proponents of new automobile technology may not have considered the affordability and rate of uptake of such vehicles among the Australian urban population. Nor has there been significant broad discussion of the social distribution of current vehicle types or the rate of vehicle turnover among socioeconomic categories. Questions surround the spatial distribution of the automobile fleet and whether the households that can afford new automobile technology are likely to be the most vulnerable to the effects of declining energy security and carbon emissions pricing. Will the most auto-dependent and socially vulnerable households be able to access the fuel-economy advantages of new vehicle technology? Although a technological change in the Australian motor vehicle fleet may appear increasingly feasible to policy makers, questions remain about how such a change might integrate with existing transport patterns in Australian cities.

Dependence on automobiles is highly spatially differentiated in Australian cities. Differences in household dependence on auto-

mobiles, distances traveled, and motor vehicle ownership will be critical in a time of energy constraint and carbon limits. In the absence of mitigating factors, the most car-dependent households in Australian cities will be exposed to far greater fuel costs that those with low automobile dependence. This may be acceptable on an austere user-cost or "polluter pays" basis, but patterns of automobile dependence intersect with important sociospatial patterns that will confound policy attempts to deploy new vehicle technology as a means of adapting to energy and carbon constraints, especially if conventional market processes are used as delivery mechanisms.

The differential transport patterns set out here are linked to socioeconomic differences observed in the structure of Australian cities. Such differences unevenly allocate household exposure to higher travel costs due to higher petroleum prices. The spatial distribution of various socioeconomic groups is heavily conditioned by spatial housing markets that intersect with transport systems. Dodson and Sipe demonstrated that the intersection of housing markets and transport patterns is a key factor in shaping oil vulnerability in Australian cities (*18–20*).

These regressive urban structural effects are depicted in Dodson and Sipe's analysis of oil vulnerability in Australian cities (*18*). The authors developed VIPER (vulnerability index for petroleum expenses), a method for assessing relative oil vulnerability within cities. This index combines measures of household socioeconomic status and car dependence taken from census variables into a single index of household exposure to the effects of higher fuel prices. Dodson and Sipe's analysis for Brisbane is presented in Figure 1. Households in the inner and middle suburban zones, especially within the central business district (CBD) and immediately adjacent areas to the west and the north exhibit low levels of exposure to declining energy security and rising fuel prices. Beyond these areas, a wide zone of moderate oil vulnerability covers many of Brisbane's middle suburban areas. In contrast, the areas of highest oil vulnerability are found among Brisbane's outer northern, outer western, and southeastern suburbs, in particular the growth zones toward Caboolture, Ipswich, and Beenleigh.

Dodson and Sipe's analysis shows that oil vulnerability is a socially regressive condition—the relatively weaker socioeconomic households in Brisbane's outer suburbs (and in comparable outer suburban zones in other Australian cities) are the most vulnerable to the adverse consequences of higher global petroleum prices due to declining energy security or carbon pricing (*18*). This reflects the urban structure of Australian cities, in which wealthier and higher-income groups tend to be located in inner high-value housing locations, whereas lower-income households are more likely to be in lower-value housing markets in outer and fringe locations. This structure contrasts with that typically found in U.S. cities, where middle and outer suburbs are often sites of relatively higher income and wealth compared with inner urban zones.

The weaker socioeconomic status of car-dependent outer suburban households in Brisbane will also impede their financial capacity to adapt to rising transport fuel costs through purchase of an alternative energy vehicle. But the problems are more complex because vehicle ownership and household social status are not the only factors shaping household vehicle adaptability. Because private motor vehicles are a major expenditure and a sunk cost for households, the characteristics of the existing motor vehicle fleet will affect household adaptability. Lower-income households may be trapped by ownership of old, inefficient vehicles, unable to afford the new vehicle technology.

ANALYZING THE URBAN AUTOMOBILE FLEET

Social analyses of the composition of the private motor vehicle fleet are rare in Australia and uncommon elsewhere. The questions raised earlier suggest that understanding the age, size, and distribution of private automobiles will be an important dimension of any improvements to the environmental efficiency of urban transport systems based on changes in vehicular and fuel technology. Registration of private motor vehicles can offer insight into these issues, especially when combined with social variables from other sources, such as the census. The remainder of this paper presents the results of an analysis of Brisbane's motor vehicle fleet. Those results are then linked to social data to develop an improved understanding of the relationship between sociospatial oil vulnerability and motor vehicle ownership. The analysis was guided by three questions: What differences can be observed in the spatial distribution of vehicle characteristics within Brisbane? What are the social implications of this vehicle distribution? What implications do spatial patterns of vehicle distribution have for improved resource and environmental efficiency in Brisbane?

Data on motor vehicle registrations collected by the Queensland state government motor vehicle registry, structured spatially at the postcode level, were obtained for the analysis. The data set from the fourth quarter of 2008, which contained 441,930 vehicle records, was used. Data were analyzed for the Brisbane urban area expanded to accommodate post codes, following Dodson and Sipe (*18*), who used urban areas for Australian cities as the spatial category for analysis of oil vulnerability, at the smallest census collection district scale. The use of postcodes as the geographic unit in the analysis allows comparison of vehicle fleet and social variables along a number of dimensions. The investigated dimensions include engine size, vehicle ownership, and vehicle age.

Automobile Engine Size

Engine size is an important determinant in vehicle fuel economy with the number of engine cylinders providing a basic ordinal indicator of relative fuel consumption. In Brisbane, the private automobile fleet is dominated by cars having four-cylinder engines. Of the 441,914 private motor cars registered in Brisbane in 2008, 62% had four-cylinder engines and 30.9% had six-cylinder engines. Eight-cylinder vehicles made up only 5% of the private car fleet. Vehicles with other quantities of cylinders made up just 2.1% of the total car fleet and are antique, exotic, or custom vehicles.

A total of 36.8% of the Brisbane automobile fleet has engines of more than four cylinders. Given that six-cylinder automobiles are indicative of higher fuel consumption, this proportion of larger engines presents policy makers with a significant challenge in shifting the fleet toward vehicles of higher fuel economy.

Vehicle Age

The age structure of the Brisbane private car fleet shows a weak bimodal clustering with a long tail. The fleet is relatively old, 12.2 years, with a mean year of manufacture of late 1996 (Figure 2). There was a sharp fall in vehicle registrations in the 2008–2009 period, attributable to a combination of high fuel prices during 2008 and economic fragility associated with the global financial crisis.

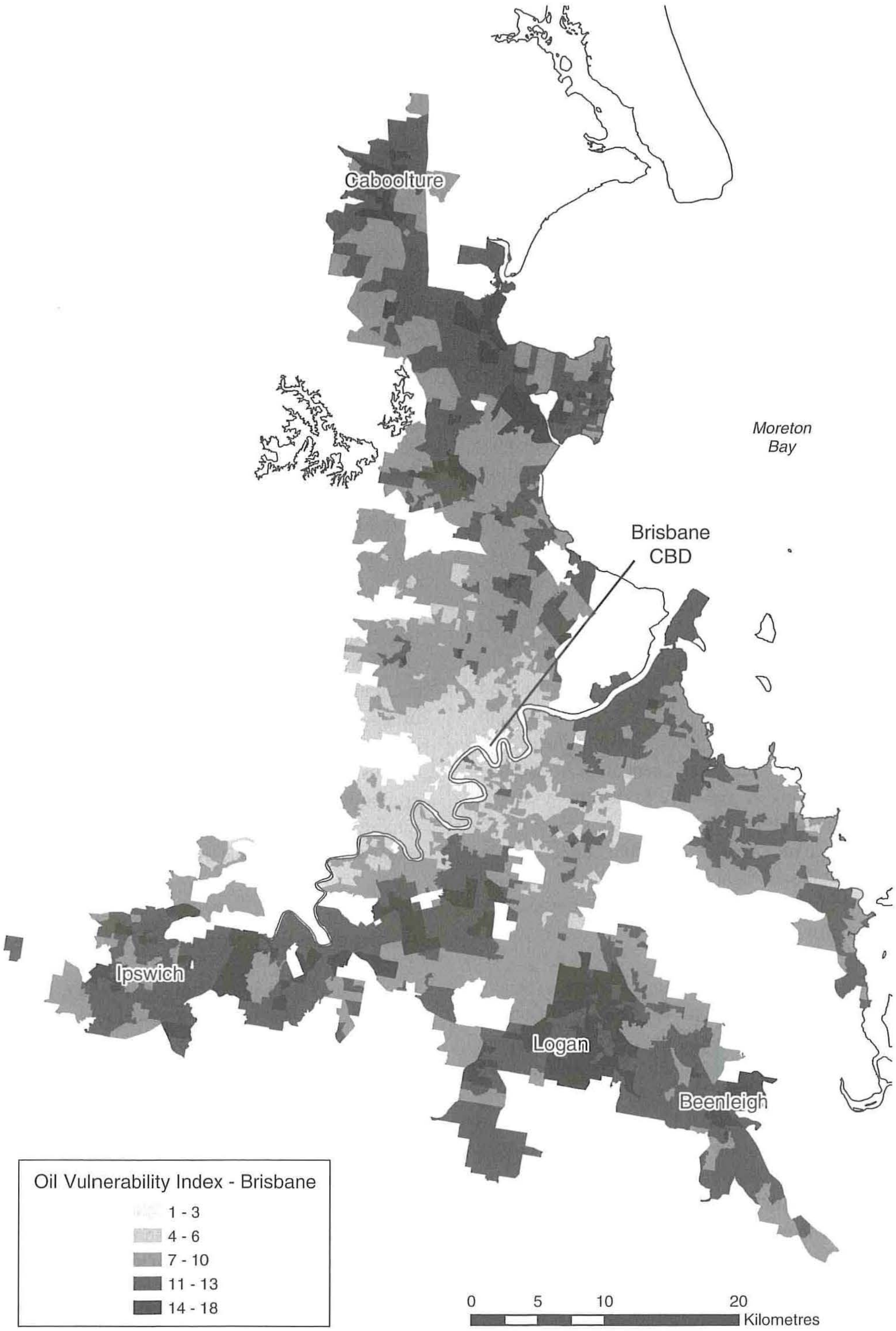


FIGURE 1 Spatial distribution of relative household oil vulnerability in Brisbane.

Such a decline in automobile registrations is both positive and negative. A decline in registrations indicates a moderating of growth in the private car fleet, meaning fewer cars in use. However, from a technology transition perspective, the drop in new auto registrations suggests a lag in the uptake of vehicles having improved fuel efficiency, with implications for improvements of overall fleet fuel economy.

Vehicles per Household

Census data were used to assess the relative numbers of automobiles per household at the postcode level. There appears to be a coarse spatial difference in proportions of cars per household across Brisbane postcodes. Households in the inner areas immediately surrounding the CBD and in the inner northern and inner southeastern areas of the city

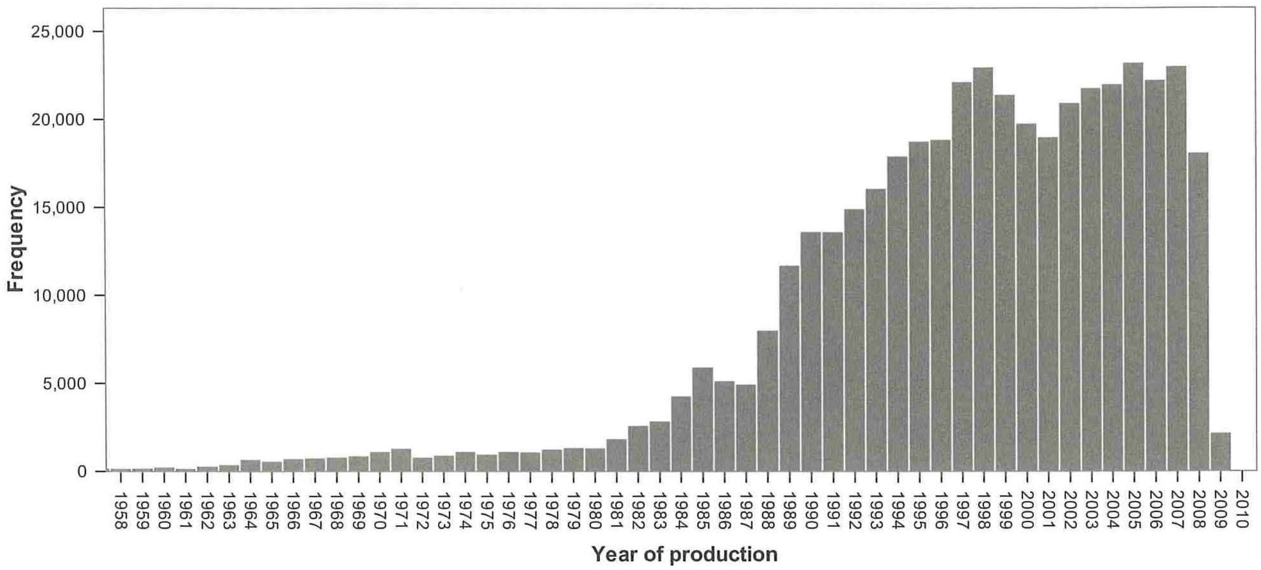


FIGURE 2 Frequency distribution of Brisbane private car fleet by year of production to March 2009 (mean = 1,996.8, SD = 9.241, N = 441,436).

exhibit relatively low levels of cars per household, reflecting the good public transport and better walkability of such zones. Some outer eastern and far outer northern postcodes also have relatively moderate numbers of cars per household. By contrast, postcodes with much higher relative levels of cars per household are located in Brisbane's northwest, west, southwest, and south, especially in the Ipswich corridor running southwest of the city center. These results demonstrate that motor vehicle ownership is highly spatially differentiated. Policies to transition automobile technology will thus affect these households to a greater extent because they have more automobiles to convert.

Distribution of Large Automobiles

Although relative number of automobiles per household in a postcode serves as a coarse indicator of the potential distribution of relative burden in any technological transformation of the Brisbane car fleet, a better indicator is large vehicles as a proportion of the fleet. This is because larger cars generally have higher emissions and therefore are reasonable targets for priority removal from the fleet or conversion to a low-carbon fuel. A cluster of postcodes with low proportions of large-engine autos is observable to the immediate west and south of the Brisbane CBD. This cluster is surrounded by a wider zone of postcodes with moderate rates of large cars in Brisbane's middle northern and middle southern suburbs. High rates of large vehicles per postcode are observed among postcodes in the northwest, west, southwest, and south of Brisbane, particularly in postcodes beyond approximately 10 km from the CBD. For example, high proportions of large cars can be observed in the Ipswich subregion, to the southwest of Brisbane.

Distribution of Old Automobiles

The relative age of the private automobile fleet is another factor that could impede the mass introduction of new vehicle technology. In general, secondhand cars tend to be cheaper to purchase than new cars and could compete with new fuel-efficient vehicles. Very low proportions of old automobiles are found within inner Brisbane postcodes, particularly those within 10 km of the CBD. In contrast, much higher proportions of old cars per household are found in the middle and outer Brisbane postcodes, particularly those in the north, southwest, and south. These patterns mean that different areas of the city face varying adjustment burdens in relation to higher fuel prices and the rollout of new vehicle technology.

Automobile Age and Engine Size

Older automobiles are likely to perform worse for fuel efficiency and carbon emissions than newer cars, given advances in technology over time, and large-engine cars typically have higher levels of fuel consumption than smaller cars of a similar age. Postcodes with higher proportions of large older autos will face greater fleet adjustment than those with much lower proportions of large old cars. The proportion of old large cars by Brisbane postcode was mapped, where old cars are those with a production year before 1996 and large cars are those with six or more cylinders.

The distribution of old large automobiles by postcode in Brisbane is highly uneven (Figure 3). In general, inner urban areas immediately surrounding the Brisbane CBD and extending to the west and east of the CBD have very low relative proportions of old large cars per household. This band is surrounded to the north and south by postcodes with relatively moderate proportions of old large cars. In contrast, postcodes to the far north, southwest, south, and southeast exhibit among the highest proportions of old large cars per household. In this analysis, the best-performing postcodes for car age and size had proportions of old large cars that were around half the level of those that performed with the highest proportions of such vehicles.

Comparing Automobile Fleet Efficiency and Socioeconomic Status

The analysis of fleet efficiency and socioeconomic status comprises two parts. First, the VIPER oil vulnerability analysis method described and deployed by Dodson and Sipe (*18*) was replicated to provide an

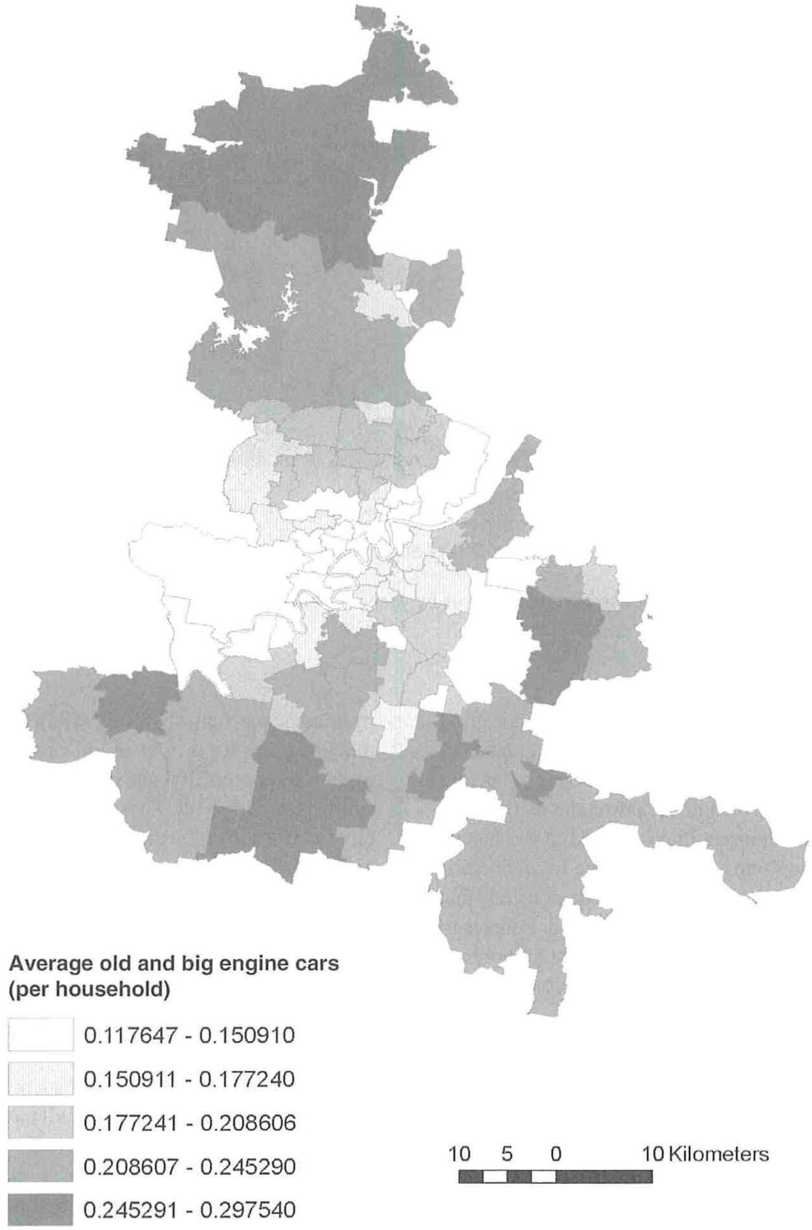


FIGURE 3 Number of old large cars per household for Brisbane postcodes.

assessment of the relative oil vulnerability of Brisbane postcodes. Two groups were constructed comprising the best 20 postcodes and worst 20 postcodes, respectively, within Brisbane by VIPER score. The construction of these two groups thus permitted a comparison of various car-fleet characteristics relative to levels of socioeconomic oil vulnerability. The results of this analysis, shown in Table 1, give insight into the relative adjustment task for highly oil vulnerable postcodes compared to postcodes with low oil vulnerability.

This analysis reveals some important differences in the composition of the vehicle fleet in low and high VIPER postcodes. The 20 least-vulnerable postcodes performed better than the 20 most-vulnerable postcodes on measures such as automobiles per household, automobile age, proportion of large-engine automobiles, proportion of old automobiles, and proportion of old large-engine automobiles. Some of these variables are particularly distinctive. Postcodes in the most-vulnerable VIPER group had median proportions of old automobiles that were almost 16 percentage points greater than those in the least-vulnerable VIPER group, whereas there was a difference of 10 percentage points in favor of the least-vulnerable VIPER group for the large automobile variable. A comparable difference of nine percentage points is observed for the least-vulnerable VIPER postcodes and most-vulnerable VIPER postcodes for the proportion of old large autos.

The second part of the analysis tests the relationship between the VIPER score and the proportion of old large automobiles in a postcode (Figure 4) by using linear regression. The analysis reveals a very strong positive relationship between VIPER score and the proportion of old large autos in a postcode. Postcodes with low VIPER scores (and hence low exposure to the socioeconomic risks of higher fuel prices) tended to have a much lower proportion of old large

TABLE 1 Old Large Cars as Proportion of Car Fleet

Variable	Statistics	Low VIPER POAs	High VIPER POAs
Proportion of total cars (per household)	Mean	0.725	0.814
	Maximum	0.972	1.006
	Minimum	0.502	0.571
	Standard deviation	0.107	0.133
	Median	0.712	0.802
Distribution of car age (years)	Mean	14.4	15.4
	Maximum	18	18
	Minimum	11	11
	Standard deviation	2.15	1.79
	Median	14	15
Proportion of big-engine cars	Mean	0.240	0.311
	Maximum	0.362	0.388
	Minimum	0.172	0.227
	Standard deviation	0.047	0.049
	Median	0.226	0.325
Proportion of old cars	Mean	0.323	0.485
	Maximum	0.391	0.678
	Minimum	0.261	0.373
	Standard deviation	0.041	0.078
	Median	0.319	0.474
Proportion of old and big-engine cars	Mean	0.147	0.240
	Maximum	0.177	0.289
	Minimum	0.126	0.204
	Standard deviation	0.001	0.024
	Median	0.145	0.235
Proportion of new cars	Mean	0.040	0.028
	Maximum	0.054	0.042
	Minimum	0.024	0.016
	Standard deviation	0.007	0.006
	Median	0.039	0.028
Proportion of new and big-engine cars	Mean	0.011	0.005
	Maximum	0.019	0.010
	Minimum	0.006	0.003
	Standard deviation	0.003	0.002
	Median	0.011	0.005

NOTE: VIPER = vulnerability index for petroleum expenses. POA = postcode.

vehicles than did postcodes with high VIPER scores. This relationship is very robust with an R^2 value of 0.638 and a T-value of 13.5. These results suggest that attempts to overcome the impact of higher fuel prices through the use of market-based technological advances will be confounded by sociospatial patterns in Australian cities. Such a problem is likely to affect other metropolitan regions, especially the dispersed cities of North America.

CONCLUSIONS

The world faces considerable energy security and climate mitigation challenges in coming years and decades. Urban transport systems will be strongly affected by the shifting energy environment. Many government policy settings assume that advances in motor vehicle technology combined with a shift to low-carbon fuel sources will enable a relatively smooth transition to an energy-secure, greenhouse-neutral vehicle fleet. Such a policy is being pursued in Australia through the federal government's Green Vehicle Plan and associated schemes to reorient the Australian automotive sector toward the design and production of environmentally friendly vehicles. Such policies echo those found in other jurisdictions, such as the United States, where the government is undertaking a major restructuring of the domestic automotive manufacturing sector focused around energy-efficient vehicle technology.

Although new technology holds considerable appeal among politicians and policy makers, a number of problems and risks are associated with a reliance on technology as a salve for urban transport problems. The analysis of the vehicle fleet in Brisbane presented in this paper illustrates these problems in three ways. First, the age structure of the Brisbane motor vehicle fleet is such that even if all new vehicles entering the market within a few years were fully electric or petrol–electric hybrids, there would still be a long period, potentially a decade or more given current vehicle age patterns, before even half the vehicle fleet is made up of such vehicles. The recent drop in motor vehicle sales due to global financial problems poses a further setback for the acceptance of new vehicle technology, given the stockpiles of existing new vehicles.

Second, the sociospatial dimensions of the vehicle fleet mean that the oldest and largest motor vehicles are likely to be found in areas where households can less afford vehicle upgrades. These zones are also likely to experience the greatest economic stress from the global financial crisis (*21*). There is already evidence that the 2004–2008 oil-price shock was a key contributing factor to the current recession in the United States (*22*). Although there has been only modest evidence to support a similar relationship in Australia, the problem of high fuel prices and weak economic performance may limit household ability to afford new vehicle types.

Third, the relatively weaker household capacity to afford motor vehicles intersects with the structure of transport and travel behavior outlined previously. In combination, these patterns will impede objectives for reduction of carbon pollution because the areas with the highest levels of dependence on cars for transport also are typically those with relatively modest socioeconomic capacity to afford new vehicle types. Households of lower socioeconomic status therefore could continue to drive longer distances in old large cars than will more prosperous households located more centrally. The result is a distinct friction in any process of transforming the composition of the urban vehicle fleet to target the oldest, largest vehicles that are also used the most.

Efforts to reduce urban transport dependence on increasingly insecure petroleum and to reduce the carbon intensity of urban transport through the deployment of new vehicle technologies are unlikely to offer a comprehensive transformation of urban transport patterns, especially in Australia. In the United States, the car allowance rebate offered a vehicle trade-in rebate of up to $4,500 for inefficient vehicles conditional on purchase of a new, efficient vehicle. Such a policy has been mooted for Australia and are likely to be attractive only to a subrange of owners of inefficient vehicles who can afford the purchase costs of a new vehicle beyond the $4,500 trade-in rebate, and it will likely have only a modest effect on fleet composition, at a $1 billion cost.

A much wider strategic approach to coping with declining energy security and to reducing carbon emissions will be needed, in which technology takes a less prominent role and other methods, such as the rollout of improved public transport networks and associated coordinating institutions (*16, 19*) are given greater emphasis. Although

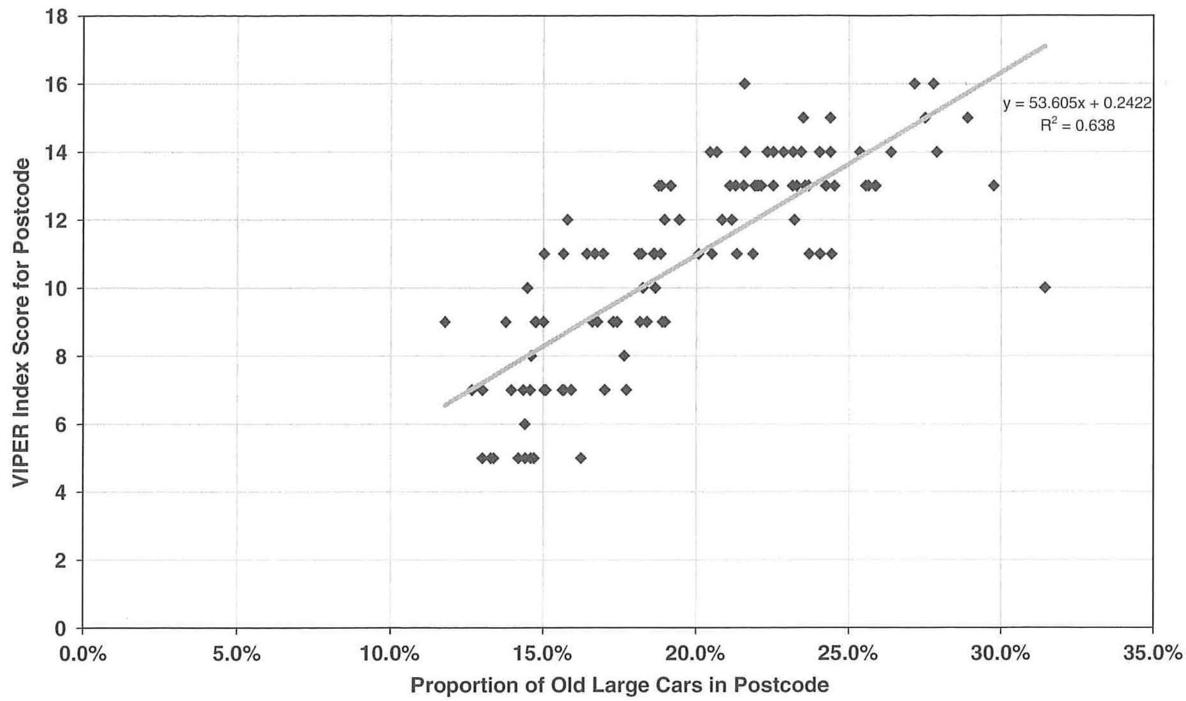


FIGURE 4 Proportion of old large cars versus VIPER score for Brisbane postcodes.

new private vehicle and fuel technologies will provide an element within urban transport futures, they probably will be only a component rather than a dominant mode. Policies that fail to recognize this wider context will be insufficient to offer dispersed car-dependent cities secure low-carbon transport. Such wider, more nuanced policy thinking will be necessary if North American and Australian cities are to maintain high levels of mobility under declining petroleum security and increasing carbon restraint.

Further work is needed in this area of research and analysis. The authors will further the investigation of vehicle fleet composition and household social status by adding data on fuel consumption by vehicle type and data on vehicle kilometers traveled at the postcode level, which will provide a richer depiction not only of the vehicle fleet but also of the relative levels of fuel and carbon consumption of that fleet under current conditions.

REFERENCES

1. Carr, K., and K. Rudd. *A New Car Plan for a Greener Future.* Minister for Innovation, Industry, Science and Research, Canberra, Australia, 2008.
2. Moriarty, P., and D. Honnery. What Energy Levels Can the Earth Sustain? *Energy Policy,* Vol. 37, No. 7, 2009, pp. 2469–2474.
3. *World Energy Outlook 2007: Focus on China and India.* International Energy Agency and Organisation for Economic Cooperation and Development, Paris, 2007.
4. Yi-Chong, X. China's Energy Security. In *Energy Security in Asia* (M. Wesley, ed.), Routledge, New York, 2007.
5. Kleveman, L. *The New Great Game: Blood and Oil in Central Asia.* Atlantic Monthly Press, New York, 2003.
6. Klare, M. *Rising Powers, Shrinking Planet: The New Geopolitics of Energy.* Metropolitan Books, New York, 2008.
7. Campbell, C., and J. H. Laherrere. The End of Cheap Oil. *Scientific American,* Vol. 278, No. 3, 1998, pp. 78–83.
8. Deffeyes, K. *Hubbert's Peak: The Impending World Oil Shortage.* Princeton University Press, Princeton, N.J., 2001.
9. Campbell, C. *Oil Crisis.* Multi-Science Publishing, London, 2005.
10. Deffeyes, K. *Beyond Oil: The View from Hubbert's Peak.* Farrar, Strauss and Giroux, New York, 2005.
11. Syed, A., R. Wilson, S. Sandu, C. Cuevas-Cubria, and A. Clarke. *Australian Energy: National and State Projections to 2029–2030.* Australian Bureau of Agriculture and Regional Economics, Canberra, 2007.
12. *Energy in Australia.* Department of Resources Energy and Tourism, Canberra, Australia, 2008.
13. Intergovernmental Panel on Climate Change. *Climate Change 2007: Mitigation.* Cambridge University Press, Cambridge, United Kingdom, 2007.
14. *Garnaut Climate Change Review: Final Report.* Garnaut Review of Climate Change, Melbourne, Australia, 2008.
15. Bracks, S. *Review of Australia's Automotive Industry.* Australian Government, Canberra, 2008.
16. Mees, P. *Rethinking Public Transport in Sydney.* UFP Issues Paper 5. Urban Frontiers Program, University of Western Sydney, Australia, 2000.
17. Sivak, M., and O. Tsimhoni. Fuel Efficiency of Vehicles on US Roads: 1923–2006. *Energy Policy,* Vol. 37, No. 8, 2009, pp. 3168–3170.
18. Dodson, J., and N. Sipe. Oil Vulnerability in the Australian City: Assessing Socioeconomic Risks from Higher Urban Fuel Prices. *Urban Studies,* Vol. 44, No. 1, March 2007, pp. 37–62.
19. Dodson, J., and N. Sipe. *Shocking the Suburbs: Oil Vulnerability in the Australian City.* UNSW Press, Sydney, Australia, 2008.
20. Dodson, J., and N. Sipe. Shocking the Suburbs: Urban Location, Homeownership and Oil Vulnerability in the Australian City. *Housing Studies,* Vol. 23, No. 3, 2008, pp. 377–401.
21. Baum, S., and W. Mitchell. *Red-Alert Suburbs: An Employment Vulnerability Index for Australia's Metropolitan Regions.* Centre of Full Employment and Equity, University of Newcastle, and Urban Research Program, Griffith University, Brisbane, Australia, 2009.
22. Hamilton, J. *Causes and Consequences of the Oil Shock of 2007–08.* Brookings Institution, Washington, D.C., 2009.

The Travel Analysis Methods Section peer-reviewed this paper.

Two-Stage Hedonic Price Model for Light-Duty Vehicles

Consumer Valuations of Automotive Fuel Economy in Maine

Qin Fan and Jonathan Rubin

Consumers' marginal willingness to pay for a unit change of automotive fuel economy was estimated through development of a hedonic regression of new automobiles sales. The research combined national data on vehicle attributes with a unique data set that contains demographic information on all new vehicles registered in Maine in 2007. The research estimates the impact of demographic factors on consumer demands for fuel economy by generating a function for fuel economy demand in a second-stage hedonic model. Results show that consumers undervalue the long-run fuel savings of vehicle ownership, but they significantly value short-run fuel savings. Age and education are positively correlated with fuel economy demand, whereas income is statistically insignificant. Car consumers' net benefits from an increase in fuel economy from 25 to 35 mpg are computed from the fuel economy demand curve and are approximately \$2,232. Strengthening corporate average fuel economy standards is reasonable because consumers can receive significant net benefits from increasing fuel economy.

A growing recognition of climate change, energy security, and volatile fuel prices leads to analyses of policy actions intended to reduce the fuel consumption of light-duty vehicles and reduce greenhouse gas (GHG) emissions in the United States. Policy instruments include information programs, federal corporate average fuel economy (CAFE) standards, economic incentives for certain highly fuel efficient vehicles, and subsidies for research in new vehicle technologies and fuels. The cost-effectiveness of these policies is in turn strongly dependent on consumer valuations of automotive fuel economy.

The focus of this study is how individuals value not only incremental changes in fuel economy but also larger changes—on the order of 40%—as are mandated under the Energy Independence and Security Act of 2007 (*1*). Moreover, the cost-effectiveness of these policies is contingent on the characteristics of each consumer, such as education level, age, and income. Estimation of the impact of demographic and financial factors on consumer purchasing decisions is another purpose of this paper.

CAFE standards, established by the U.S. Energy Policy and Conservation Act of 1975, specify minimum fuel economy standards of 27.5 miles per gallon (mpg) for the entire fleet of passenger cars and of 22.2 mpg for light-duty trucks (*2*). Effective in 2011, the proposed reformed size-based CAFE standards regulate each individual vehicle with a specific fuel economy target based on a vehicle's footprint (a product of multiplying a vehicle's wheelbase by its track width). The reformed CAFE standards are proposed to achieve a combined fuel economy of 35 mpg for the entire fleet of both passenger cars and light-duty trucks manufactured for sale in 2020 (*3*).

School of Economics, University of Maine, 5782 Winslow Hall, Orono, ME 04469. Corresponding author: Q. Fan, qin.fan@umit.maine.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 119–128.
DOI: 10.3141/2157-15

LITERATURE REVIEW

Previous studies have estimated consumer willingness to pay (WTP) for automotive attributes by using a single-stage hedonic price model (*4–7*); others have estimated consumer preferences for vehicle attributes by using the discrete choice model (*8–10*). Although surveys have the advantage of providing additional information on attitudes and motivations for purchasing behavior, discrete choice models are confined to a few hundred observations from surveys of consumers (*11*). Hedonic models have the advantage of using detailed data from thousands of observations of automobile prices and attributes.

Although the hedonic price model was developed for automobiles early in 1939 (*12*), the attribute of fuel economy was not included in hedonic regressions until the 1980s (*5, 13, 14*). Since then, a number of studies have estimated consumer valuations of automotive fuel economy by using hedonic price approaches (*7, 15, 16*). Many of these studies, however, used a single-stage model and were based on dated information and databases.

Unlike the previous literature on the hedonic value of fuel economy, this paper uses a two-stage model to estimate a demand curve for fuel economy. A two-stage model not only estimates the equilibrium value, as is done in a single-stage analysis, but also estimates demographic impact on fuel economy demands. The second-stage analysis, however, presents a number of problems that researchers continue to address. In particular, there is a challenge of dealing with endogeneity (*17–19*), leading to the failure of ordinary-least-squares estimation (*20*). The correlation between implicit prices of attributes and the error term can be solved in an ideal market where consumer taste for characteristics of a good stay the same. Unfortunately, there is no such market in real life unless one can utilize abundant databases that would reduce the taste difference in diverse markets (*19*). Beron et al. conducted two-stage and three-stage least-squares regressions to overcome the endogeneity problem by using instrumental variables (*20*); however, appropriate instruments that are exogenous and correlated with the endogenous variables, but uncorrelated with error terms, are difficult to obtain. The lack of sufficient instruments might lead to the under-identification problem. The following models give a good, but not ideal, methodology for solving the problems presented. The results are consistent with theoretical expectations.

THEORETICAL ANALYSIS

A first-stage hedonic model typically estimates implicit prices P_i of component characteristics of goods that are represented by Z_i. In the automobile market, the utility of an automobile depends on consumer valuations of all attributes. According to Rosen's study on the hedonic model (*21*), the partial derivative of the price of an automobile with respect to a certain attribute represents consumers' optimal bids for a marginal increase of that attribute in the demand market, and it represents the amount of money producers are willing to accept for increasing a unit characteristic in the supply market where the market equilibrium is achieved. A first-stage hedonic function is generally considered to be nonlinear since consumers' WTP and producers' willingness to accept a marginal shift of a characteristic change as Z_i changes.

Pickering et al. suggested the principal vehicle attributes that consumers are concerned about: safety, vehicle class, horsepower, weight, acceleration, fuel economy, drive type (e.g., all-wheel drive), body type, transmission, and manufacturer (*13*). The hedonic price for each characteristic is derived from Equation 1:

$$P(Z_k) = \frac{\partial P}{\partial Z} = P_k(Z_1, Z_2, Z_3, \ldots, Z_n) \tag{1}$$

where $P(Z_k)$ equals the implicit price of specific attribute Z_k and Z equals different attributes.

The demand function in a second-stage hedonic model includes the substitute and complement attributes of a particular characteristic (*17*). Income and other demographic variables are also important factors that could affect consumer demand for different characteristics. The theoretical model is shown in Equation 2:

$$D = f(P_i, P_s, P_c, Y, \tau) \tag{2}$$

where

D = consumer demands for a specific characteristic,
P_i = price of the characteristic of interest,
P_s = price of substitutes,
P_c = price of complements,
Y = income, and
τ = demographic variables, such as age, gender, and education.

The focus of a second-stage hedonic model is to measure consumer net benefits by using a compensated demand function. Consumer net benefits from a change in the quantity of an attribute are simply the integral of the compensated demand function given a certain quantity change of vehicle attributes (*17*).

DATA AND METHODOLOGY

The database consists of two parts: one for passenger cars and one for light-duty trucks. They are both constructed from comprehensive data sets containing detailed vehicle attributes, town-level household income, and town-level education indicators. Town-level income and education are used to represent each consumer who purchases a specific new vehicle of model year (MY) 2007 in Maine. Age of an individual is calculated by the date of birth provided in vehicle registration data. National vehicle attributes and sales data of MY 2007 are from Ward's Automotive Group (*22*). Sales data are merged into vehicle-attributes data by vehicle trim level. Maine vehicle registration records were purchased through Information Resource of Maine and the Maine Bureau of Motor Vehicles (*23, 24*). Maine demographic data were obtained from the U.S. Census Bureau. The demographic information, including income and education, is then added to Maine vehicle registration data by town. The car registration database with complete vehicle attributes and demographic information has 523 observations, and the truck registration database with complete truck attributes and demographic information has 2,100 observations. Table 1 give the summary statistics of the final data.

EMPIRICAL MODEL SPECIFICATION

The Box–Cox transformation methodology was used to determine two parameters, θ and λ, in the quadratic functional form in Equation 3 (*17*). Both parameters are found to be zero, indicating that log-log functional form best fits the data.

$$P^{(\theta)} = \alpha + \sum_{i=1}^{N} \beta_i Z_i^{(\lambda)} + 0.5 \sum_{i=1}^{N} \sum_{j=1}^{N} \delta_{ij} Z_i^{(\lambda)} Z_j^{(\lambda)} \tag{3}$$

where

$$P(Z)^{(\theta)} = \frac{\left[\left(P(Z)^{(\theta)} - 1\right)\right]}{\theta}$$

$$Z^{(\lambda)} = \frac{\left(Z^{(\lambda)} - 1\right)}{\lambda}$$

P = price vector,
Z = attribute,
$P^{(\theta)} = (P^{(\theta)} - 1) / \theta$ if $\theta \neq 0$
$= \ln(P)$ if $\theta = 0$

The same transformation applied to λ where λ is an exponent of Z and θ is an exponent of P.

Therefore, the empirical model in the first-stage hedonic study is as shown in Equation 4:

$$\begin{aligned}\log(\text{MSRP}) = {} & \beta_1 + \beta_2 \text{class} + \beta_3 \log(\text{MPG}) + \beta_4 \log(\text{weight}) \\ & + \beta_5 \log(\text{HP.WT}) + \beta_6 \text{transmission} \\ & + \beta_7 \text{manufacturer} + e_i \end{aligned} \tag{4}$$

where class is a vector of dummy variables. Other variables are listed in Table 1. In the car data set, class = 1, 2, 3, 4, which indicates four dummy variables—luxury, large, mid, and small—respectively, representing luxury cars, large cars, midsize cars, and small cars. Small is omitted as the reference. In the truck data set, class = 1, 2, 3, which indicates SUV, van, and pickup, respectively, representing sport utility vehicles, vans, and pickup trucks. Pickup is left out as the reference. Other variables are listed in Table 1.

To examine further the differences of WTP for fuel economy across different vehicle classes, a vector of dummy variable class is introduced as a slope shifter of MPG, weight, and HP.WT in the model. Empirical models are shown in Equations 5 and 6, which represent the model for cars and trucks, respectively:

$$\begin{aligned}\log(\text{MSRP}) = {} & \beta_1 + \beta_2 \text{mid} + \beta_3 \text{large} + \beta_4 \text{luxury} + \beta_5 \log(\text{MPG}) \\ & + \beta_6 (\text{mid} \times \log(\text{MPG})) + \beta_7 (\text{luxury} \times \log(\text{MPG})) \\ & + \beta_8 (\text{large} \times \log(\text{MPG})) + \beta_9 \log(\text{weight}) \end{aligned}$$

TABLE 1 Summary Statistics for Maine Cars and Trucks, MY 2007

Type	Description	Mean	Min.	Max.
Cars				
MSRP ($)	Manufacturers' suggested retail price	30,176	9,995	182,275
small	Small cars	0.3308	0	1
mid	Midsize cars	0.2237	0	1
large	Large cars	0.09178	0	1
luxury	Luxury cars	0.3537	0	1
HP.WT(hp/lbs)	Power–weight ratio	0.06375	0.0259	0.1612
MPG	Combined fuel economy (mpg) 1/(0.55/mpg city + 0.45/mpg hway)	24.57	15.43	55.59
weight	Vehicle curb weight (lbs)	3,401	2,293	4,985
manufacturer	Automobile manufacturers containing 17 companies	7.18	1	17
transmission	Transmission, auto = 1, manual = 0	0.4629	0	1
age	Age of automobile consumer	51.21	19	106
education	Percentage of bachelor's degree or higher (%)	26.63	4.9	58.7
income	Household income	49,823	25,871	87,126
	Observations: 523			
Trucks				
MSRP ($)	Manufacturers' suggested retail price	29,737	14,497	110,822
SUV	Sport utility vehicles	0.2086	0	1
pickup	Pickup trucks	0.7376	0	1
van	Vans	0.05381	0	1
HP.WT(hp/lbs)	Power–weight ratio	0.0554	0.02712	0.09876
MPG	Combined fuel economy (mpg)1/(0.55/mpg city + 0.45/mpg hway)	18.02	13.19	33.56
manufacturer	Automobile manufacturers containing 17 companies	7.18	1	17
transmission	Transmission, auto = 1, manual = 0	0.7305	0	1
age	Age of automobile consumer	50.98	19	100
education	Percentage of bachelor's degree or higher (%)	22.74	1.4	58.7
income	Household income	463,38	18,814	87,126
	Observations: 2,100			

$$
\begin{aligned}
&+\beta_{10}(\text{mid}\times\log(\text{weight}))+\beta_{11}(\text{luxury}\times\log(\text{weight}))\\
&+\beta_{12}(\text{large}\times\log(\text{weight}))+\beta_{13}\log(\text{HP.WT})\\
&+\beta_{14}(\text{mid}\times\log(\text{HP.WT}))+\beta_{15}(\text{luxury}\times\log(\text{HP.WT}))\\
&+\beta_{16}(\text{large}\times\log(\text{HP.WT}))+\beta_{17}\text{transmission}\\
&+\beta_{18}(\text{manufacturer})+e_t \qquad (5)
\end{aligned}
$$

$$
\begin{aligned}
\log(\text{MSRP}) = {}&\alpha_1+\alpha_2\text{SUV}+\alpha_3\text{pickup}+\alpha_4\log(\text{MPG})\\
&+\alpha_5(\text{SUV}\times\log(\text{MPG}))+\alpha_6(\text{pickup}\times\log(\text{MPG}))\\
&+\alpha_7\log(\text{weight})+\alpha_8(\text{SUV}\times\log(\text{weight}))\\
&+\alpha_9(\text{pickup}\times\log(\text{weight}))+\alpha_{10}\log(\text{HP.WT})\\
&+\alpha_{11}(\text{SUV}\times\log(\text{HP.WT}))\\
&+\alpha_{12}(\text{pickup}\times\log(\text{HP.WT}))+\alpha_{13}\text{transmission}\\
&+\alpha_{14}(\text{manufacturer})+e_t \qquad (6)
\end{aligned}
$$

All variables are listed in Table 1.

For examination of the impact of demographic factors on fuel economy, three demand functions were generated for three vehicle attributes—automotive fuel economy, vehicle weight, and power–weight ratio—in a second-stage hedonic model. Demographic variables (income, age, and education) are included in each equation. It is assumed that unobservable fuel prices are included in error terms. These error terms might be correlated with each other and with endogenous variables. Seemingly unrelated regression (SUR) is used to adjust the correlated error terms across these three equations: Equation 7 shows the simultaneous equations for cars, and Equation 8 indicates the simultaneous equations for light-duty trucks:

$$
\left\{
\begin{aligned}
\log(\text{MPG}) = {}&\alpha_1+\mu_1\log(\text{PMPG})+\mu_2\log(\text{income})\\
&+\mu_3\log(\text{education})+\mu_4\log(\text{age})+\mu_5\text{luxury}\\
&+\mu_6\text{small}+\mu_7\text{mid}+\mu_8\text{domestic}+e_{t1}\\
\log(\text{weight}) = {}&\alpha_2+\gamma_1\log(\text{Pweight})+\gamma_2\log(\text{income})\\
&+\gamma_3\log(\text{education})+\gamma_4\log(\text{age})+\gamma_5\text{luxury}\\
&+\gamma_6\text{small}+\gamma_7\text{mid}+\gamma_8\text{domestic}+e_{t2}\\
\log(\text{HP.WT}) = {}&\alpha_3+\eta_1\log(\text{PHP.WT})+\eta_2\log(\text{income})\\
&+\eta_3\log(\text{education})+\eta_4\log(\text{age})+\eta_5\text{luxury}\\
&+\eta_6\text{small}+\eta_7\text{mid}+\eta_8\text{domestic}+e_{t3}
\end{aligned}
\right. \qquad (7)
$$

$$\begin{cases} \log(\text{MPG}) = \alpha_1 + \mu_1 \log(\text{PMPG}) + \mu_2 \log(\text{income}) \\ \quad + \mu_3 \log(\text{education}) + \mu_4 \log(\text{age}) + \mu_5 \text{SUV} \\ \quad + \mu_6 \text{van} + \mu_7 \text{domestic} + e_{t1} \\ \log(\text{weight}) = \alpha_2 + \gamma_1 \log(\text{Pweight}) + \gamma_2 \log(\text{income}) \\ \quad + \gamma_3 \log(\text{education}) + \gamma_4 \log(\text{age}) + \gamma_5 \text{SUV} \\ \quad + \gamma_6 \text{van} + \gamma_7 \text{domestic} + e_{t2} \\ \log(\text{HP.WT}) = \alpha_3 + \eta_1 \log(\text{PHP.WT}) + \eta_2 \log(\text{income}) \\ \quad + \eta_3 \log(\text{education}) + \eta_4 \log(\text{age}) + \eta_5 \text{SUV} \\ \quad + \eta_6 \text{van} + \eta_7 \text{domestic} + e_{t3} \end{cases} \tag{8}$$

where PMPG, Pweight, and PHP.WT are implicit prices of fuel economy, vehicle weight, and acceleration from the first-stage hedonic regression; domestic is a dummy variable that represents domestic vehicles or foreign vehicles; and other variables are as stated in Equation 4.

RESULTS

The R program was used to test the assumptions of a linear regression. It was found that the errors have constant variance and are normally distributed with a mean value of zero. Tolerance and variation inflation factor were used to test multicollinearity, and it was found that there is no multicollinearity problem in the data regression. Results of a regression are shown in Table 2.

The implicit price of fuel economy—

$$\frac{\partial \text{MSRP}}{\partial \text{MPG}} = \hat{\beta}_3 \cdot \frac{\text{MSRP}}{\text{MPG}}$$

where MSRP and MPG are mean values of these two variables—is \$208 for cars and \$233 for light-duty trucks, indicating that car buyers are willing to pay \$208 and truck buyers are willing to pay \$233 for an increase in fuel economy of 1 mpg. Car consumers are willing to spend more money on luxury cars than on small cars. Consumer WTP for both midsize cars and large cars tends to be lower than that for small cars. Truck buyers are willing to pay more for SUVs and vans than for pickup trucks. This result is consistent

TABLE 2 Estimates of First-Stage Hedonic Model, Dependent Variable Log (MSRP)

Estimations for Passenger Cars			Estimations for Light-Duty Trucks		
Independent Variable	Coefficient	SE	Independent Variable	Coefficient	SE
mid	−0.10661	(−5.024)***	SUV	8.1901	(2.23)*
luxury	0.24316	(9.480)***	pickup	11.4527	(3.13)**
large	−0.2055	(−6.571)***	log(MPG)	1.1042	(4.43)***
log(MPG)	0.16934	(2.905)*	SUV*log(MPG)	−0.8746	(−3.34)***
log(weight)	1.80923	(16.347)***	pickup*log(MPG)	−1.1975	(−4.53)***
log(HP.WT)	0.61365	(13.755)***	log(weight)	2.0483	(6.19)***
transmission	0.04413	(2.578)*	SUV*log(weight)	−0.5829	(−1.71)
BMW	0.18064	(5.881)***	pickup*log(weight)	−0.8671	(−2.56)**
Chrysler	0.11001	(3.872)***	log(HP.WT)	0.2632	(2.12)*
Porsche	0.59320	(6.240)***	SUV*log(HP.WT)	0.1972	(1.50)
Ford	0.08893	(3.212)**	pickup*log(HP.WT)	0.2319	(1.80)
GM	0.02605	(1.021)	transmission	0.06351	(4.69)***
Honda	0.16432	(4.910)***	BMW	0.3396	(4.96)***
Hyundai	−0.05728	(−1.254)	Chrysler	0.09485	(6.46)***
Kia	0.07145	(1.438)	Ford	0.1367	(8.67)***
Mazda	0.07456	(1.707)	Honda	0.1173	(4.53)***
Mitsubishi	0.01930	(0.275)	Hyundai	0.04894	(1.63)
Nissan	−0.02358	(−0.748)	Kia	−0.1123	(−3.15)**
Volkswagen	0.09993	(2.826)**	Mazda	0.03645	(0.73)
Isuzu	0.21195	(5.047)***	Mitsubishi	0.07575	(2.22)*
Suzuki	0.06749	(1.315)	Nissan	0.02997	(1.65)
Toyota	0.13092	(4.646)***	Volkswagen	0.1694	(2.46)*
Volvo	0.27905	(6.654)***	Isuzu	0.2131	(5.50)***
			Suzuki	0.09756	(1.92)
			Toyota	0.1454	(8.98)***
			Volvo	0.2905	(4.26)***
			Cami	0.01186	(0.36)
			Landrover	0.3207	(4.64)***

NOTE: *t*-statistics are in parentheses. ***Significance at the 0.001 level. **Significance at the 0.01 level. *Significance at the 0.05 level. Passenger cars $R^2 = 0.9092$, observations: 523; light-duty trucks $R^2 = 0.7919$, observations: 2,100.

with most previous findings: SUVs are the most popular vehicle type. Compared to Audi, the luxury automobile manufacturer Porsche gains the highest WTP from car consumers. Compared to GM, the manufacturer BMW gains the highest WTP from truck buyers. As expected, vehicle weight and power weight ratio are positively significant, since vehicle weight and acceleration are positive indices for vehicle safety and better performance. Consumers are willing to pay more for automatic transmission than for manual transmission.

From the regression results of Equation 5, consumer WTP for a marginal change of fuel economy is calculated by

$$\frac{\hat{\beta}_5 \cdot \text{MSRP}}{\text{MPG}}$$

$$\frac{\left(\hat{\beta}_5 + \hat{\beta}_6\right) \cdot \text{MSRP}}{\text{MPG}}$$

$$\frac{\left(\hat{\beta}_5 + \hat{\beta}_7\right) \cdot \text{MSRP}}{\text{MPG}}$$

and

$$\frac{\left(\hat{\beta}_5 + \hat{\beta}_8\right) \cdot \text{MSRP}}{\text{MPG}}$$

respectively, for small, midsize, luxury, and large cars, where MSRP and MPG represent the mean values of these two variables. The computation results show that for a marginal decrease of fuel economy, luxury car buyers are willing to pay $324, and large car buyers are willing to pay $419. For a marginal increase of fuel economy, small car buyers are willing to pay $657, and midsize car purchasers are willing to pay $41. Small car buyers may have preferences for higher fuel economy. Valuations of fuel economy by buyers of luxury and large cars drag down the WTP for fuel economy improvement to the average level of $208. Although consumers of luxury and large cars may not be intentionally unwilling to pay for higher fuel economy, luxury-car consumers' preferences for luxury features and large-car buyers' primary needs for large sizes may compromise their value on fuel economy. For truck buyers, calculation results show that people who purchase vans are willing to pay $1,822 and SUV owners or operators are willing to pay $379 for a marginal increase of fuel economy, whereas pickup buyers are willing to pay only $154 for a marginal decrease of fuel economy. The valuations of fuel economy by pickup purchasers drag down the average value of WTP to $233. Van purchasers or operators would pay more for a marginal increase in fuel economy because they are often likely to be parents who care about household fuel savings. Pickup buyers may think increasing fuel economy would compromise other important preferences for vehicles, such as powerful engine and large size.

The estimated life span is 17 years for passenger cars and 16 years for light-duty trucks (*25*). For this study, it was assumed that buyers keep their new vehicles for 10 years. The 10-year forecast for fuel prices is used to calculate a vehicle's lifetime fuel savings (*25*). Annual mileage, which decreases by year, is obtained from NHTSA (*26*). These statistics are used to compute 10-year fuel savings from a unit increase of fuel economy based on Equation 9:

$$S = \sum_{i=2007}^{2017} P_i \times m_i \times \left(\frac{1}{\text{MPG}} - \frac{1}{\text{MPG}+1} \right) \tag{9}$$

where

S = lifetime fuel savings,
P_i = fuel price in the year i,
m_i = estimated annual mileages in the year i, and
MPG = represents the mean value of MY 2007 automotive fuel economy.

The results show that the calculated 10-year fuel savings for passenger cars from a unit increase of fuel economy is $823, much higher than the estimated WTP ($208) for a marginal increase of fuel economy. The calculated 10-year fuel savings for light trucks is $1,461, also much higher than the estimated value of fuel economy ($233). This indicates that consumers are willing to pay less than the calculated lifetime fuel savings, and it suggests that both car and truck consumers undervalue fuel savings in the long run. After introduction of social benefits of $284 and $265 for cars and light trucks, respectively, as based on NHTSA's regulatory impact analysis (*26*), the calculated benefits from a unit increase of fuel economy are substantially higher than the study's estimated consumer WTP for a marginal increase of fuel economy. Consumers significantly undervalue both private savings and social benefits in the long run.

Discount rates are graphed against both fuel savings and benefits, including private fuel savings and social benefits, in Figure 1. (A similar figure for light trucks is available from the authors on request.) This graph shows a significantly high discount rate, at which the estimated value of fuel economy is equal to the calculated lifetime fuel savings. Car consumers discount future fuel savings at 44%, and truck buyers discount future fuel savings at 82%. Compared to previous studies that estimate consumer valuations of fuel economy at the national level, car consumers in Maine have much higher discount rates (*27*). This indicates that Maine automobile buyers significantly undervalue lifetime fuel savings. After introduction of social benefits, the calculated value of benefits from fuel economy improvement is much higher than the estimated value of fuel economy. The benefits are not only to private agents but to society. Both discounted fuel savings, at the discount rate of 7%, and undiscounted fuel savings are plotted against year (*9*). The graph for cars is shown in Figure 2. (A similar figure for trucks is available from the authors on request.) At the discount rate of 7%, car consumers value fuel savings for the first 3 years of car ownership, whereas truck buyers value fuel savings only for the first year of truck ownership. These results are consistent with Parry and Fischer's findings, which suggest a 3-year valuation of fuel economy (*28*). These results indicate that CAFE standards can significantly increase consumers' net benefits to the extent that consumers value fuel savings only in the short run (*29*).

With the parameter estimates as the mean values, and with standard deviation from the first-stage estimation results, 523 random variables of $\hat{\beta}_3$, $\hat{\beta}_4$, and $\hat{\beta}_5$ were then generated in the car data set and 2,100 random variables of $\hat{\beta}_3$, $\hat{\beta}_4$, and $\hat{\beta}_5$ in the truck data set. These numbers match the sample size. This additional variability (and the log-log functional form) helps identify the second-stage parameters. The varied hedonic prices are thus generated from Equation 10:

$$\begin{aligned} P_{\text{MPG}_i} &= \hat{\beta}_{3i} \times \frac{\text{MSRP}_i}{\text{MPG}_i} \\ P_{\text{weight}_i} &= \hat{\beta}_{4i} \times \frac{\text{MSRP}_i}{\text{weight}_i} \\ P_{\text{HP.WT}_i} &= \hat{\beta}_{5i} \times \frac{\text{MSRP}_i}{\text{HP.WT}_i} \end{aligned} \tag{10}$$

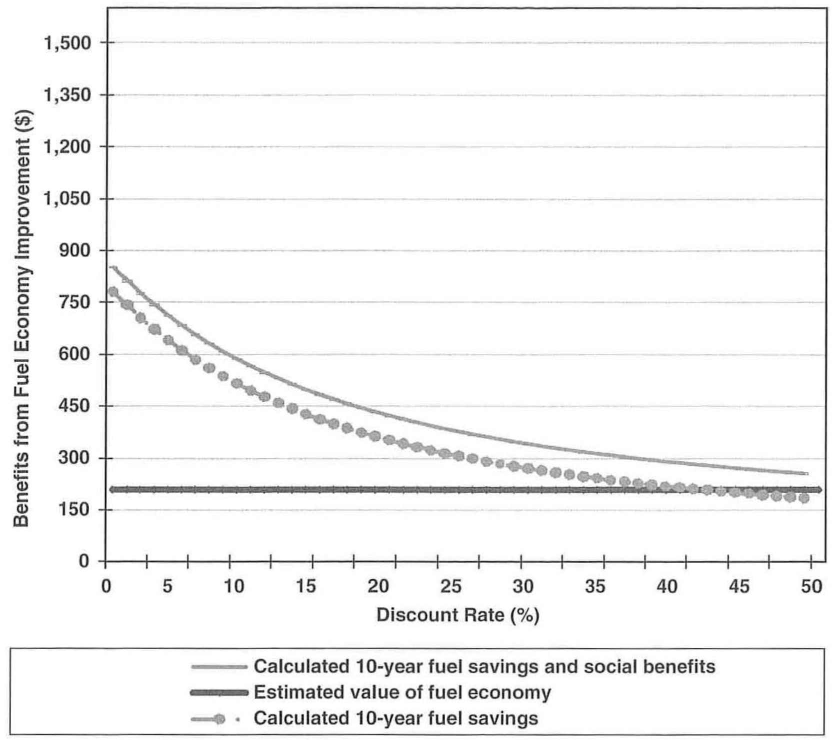


FIGURE 1 Fuel savings and social benefits from unit increase of fuel economy and discount rate (cars).

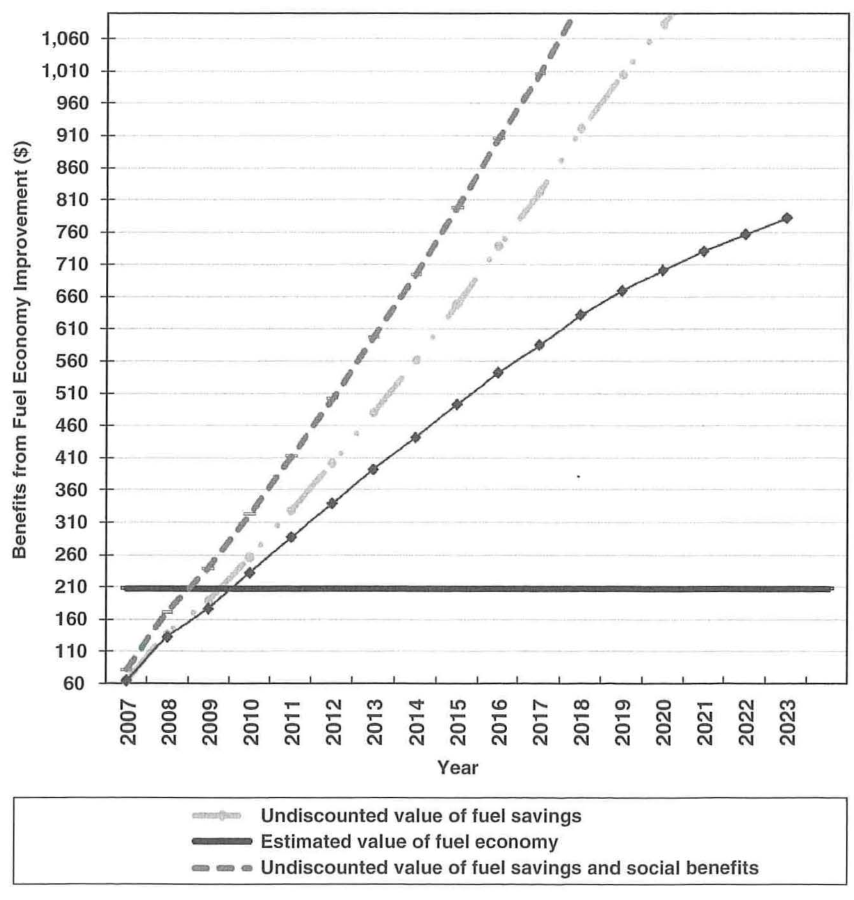


FIGURE 2 Yearly cumulative fuel savings and benefits from unit increase of fuel economy (cars).

The coefficients of the implicit PMPG have expected negative signs. The demand for fuel economy declines as the price of fuel economy increases. Income is not significant. Age in car regression results is positively significant, which indicates that the demand for fuel economy increases as age increases. Compared to large-car consumers, buyers of luxury, midsize, and small cars have a higher demand for fuel economy. In the estimations for trucks, van and SUV buyers have higher demands for fuel economy compared to pickup purchasers. Fuel economy demand of domestic car buyers is lower than that of foreign car buyers, whereas the reverse holds true for truck buyers. Estimations of results for cars and trucks are shown in Table 3.

The relationship between the implicit price of fuel economy and fuel economy itself is examined by generating the demand function of fuel economy. This relationship is an important index for the determination of desirable level of CAFE standards. Fuel economy demand curves are represented in Figures 3 and 4.

The current CAFE standard for the fleet of passenger cars is 27.5 mpg and for the fleet of light-duty trucks is 22.2 mpg. The graphs of the demand functions show that consumer demands for fuel economy approximately follow these levels. There are yet some potential benefits to an increase in the current standards of 27.5 mpg to a higher fuel economy of 35 mpg in the car graph. To the left of 35 mpg, the demand curve is slightly steeper, indicating that people are willing to pay substantial amounts for fuel economy improvement. To the right of 35 mpg, the demand curve is relatively flat, indicating that people are willing to pay relatively less for an additional unit of fuel economy. Compared to the demand curve of fuel economy for passenger cars, demand curve of fuel economy for light trucks indicates that the current CAFE standard of 22.2 mpg for the fleet of light trucks is reasonable. This occurs because the value of 22.2 mpg tends to be the turning point, after which truck buyers are willing to pay relatively less for a marginal increase of fuel economy.

The area below the demand curve represents consumers' net benefits from changing fuel economy, that is, the area is an economic measure of satisfaction with changes in fuel economy. For example, if a consumer owns a car with a fuel economy of 25 mpg, increasing fuel economy to 35 mpg equals a benefit of approximately $2,232 (the area below the demand curve and between 25

TABLE 3 Estimates of Second-Stage Hedonic Model

Estimations for Passenger Cars		
Independent Variable	Coefficient	SE
Dependent Variable: Log(MPG)		
log(PMPG)	−0.2709	(−28.79)***
log(income)	−0.02081	(−1.07)
log(education)	0.01621	(2.79)**
log(age)	0.02114	(2.84)**
luxury	0.09069	(5.19)***
small	0.05409	(3.06)***
mid	0.02513	(1.42)
domestic	−0.03642	(−3.47)***
Dependent Variable: Log(weight)		
log(Pweight)	0.2457	(13.13)***
log(income)	−0.01463	(−0.68)
log(education)	0.008075	(0.81)
log(age)	0.01357	(1.06)
luxury	−0.1461	(−7.14)***
small	−0.2346	(−12.45)***
mid	−0.1316	(−6.75)***
domestic	−0.01448	(−1.24)
Dependent Variable: Log(HP.WT)		
log(PHP.WT)	−0.04412	(−3.98)***
log(income)	0.1220	(2.42)**
log(education)	−0.01980	(−0.84)
log(age)	−0.00880	(−0.29)
luxury	0.1537	(3.36)***
small	−0.1141	(−2.52)**
mid	−0.03063	(−0.67)
domestic	0.07275	(2.59)***

Estimations for Light-Duty Trucks		
Independent Variable	Coefficient	SE
Dependent Variable: Log(MPG)		
log(PMPG)	−0.2936	(−38.15)***
log(income)	−0.00024	(−0.02)
log(education)	0.016339	(2.31)*
log(age)	0.02010	(2.39)*
SUV	0.11148	(18.03)***
van	0.1373	(8.93)***
domestic	0.01374	(2.16)**
Dependent Variable: Log(weight)		
log(Pweight)	0.3325	(10.90)***
log(income)	0.08041	(2.87)**
log(education)	−0.01450	(−1.10)
log(age)	−0.03681	(−2.36)**
SUV	−0.1031	(−7.79)***
van	−0.01756	(−0.61)
domestic	0.1298	(11.17)***
Dependent Variable: Log(HP.WT)		
log(PHP.WT)	−0.2272	(−9.81)***
log(income)	−0.02712	(−0.89)
log(education)	0.02731	(1.93)*
log(age)	−0.02119	(−1.26)
SUV	0.07659	(5.90)***
van	0.03189	(1.03)
domestic	0.07906	(6.27)***

NOTE: *t*-statistics are in parentheses. ***Significance at the 0.001 level. **Significance at the 0.01 level. *Significance at the 0.05 level. Passenger cars system weighted $R^2 = 0.6045$, observations: 523; light-duty trucks system weighted $R^2 = 0.5062$, observations: 2,100.

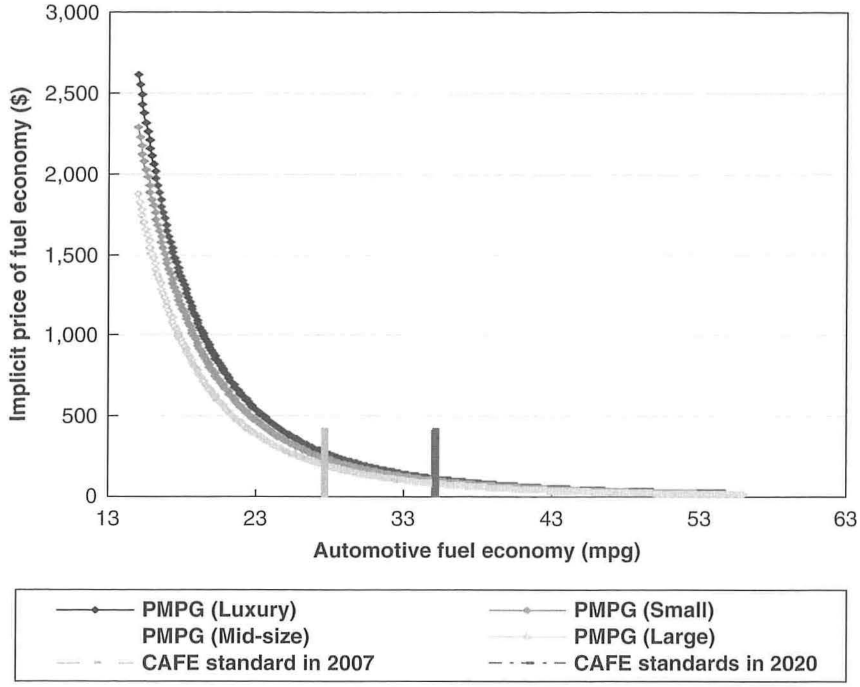


FIGURE 3 Demand curve for automotive fuel economy for cars.

and 35 mpg). Thus, the demand function suggests that from an economic perspective, it is reasonable to set a fuel economy target that is higher than the current 27.5 mpg for passenger cars, since consumers can receive significant net benefits. However, for light-duty trucks, the level of 22.2 mpg is an appropriate fuel economy standard.

CONCLUSIONS

If the presented model correctly provides information from data regression, a first-stage hedonic regression shows that for a marginal increase of fuel economy, car buyers are willing to pay $208, and truck buyers are willing to pay $233. Compared to the regression results

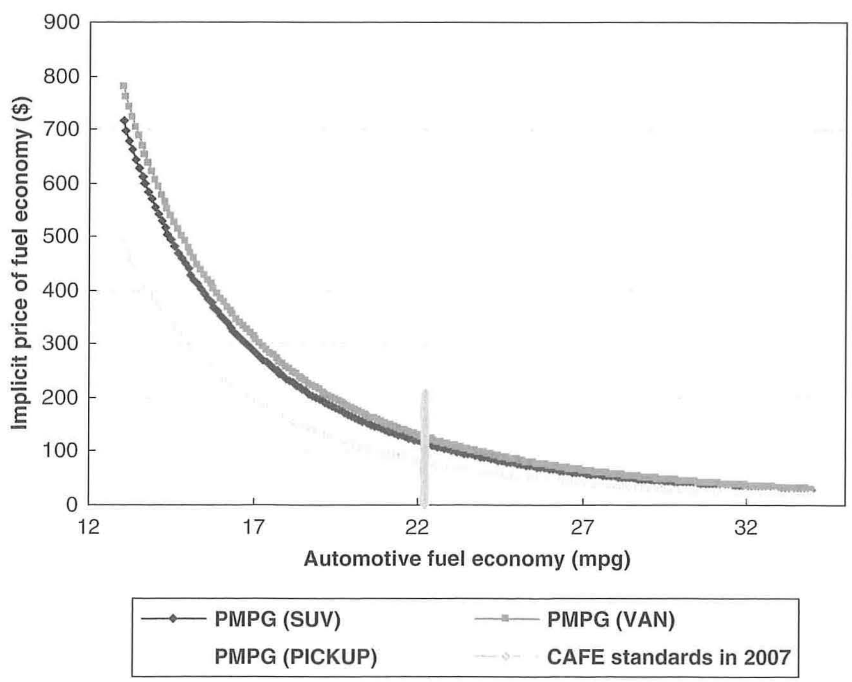


FIGURE 4 Demand curve for automotive fuel economy for light trucks.

from national data, the results from Maine state-level data indicate a much lower value of fuel economy (*6*). Moreover, consumer valuations of fuel economy differ across vehicle classes. Valuations of fuel economy by consumers of luxury and large cars drag down consumer average WTP for a marginal increase of fuel economy to $208. Luxury and large-car consumers might compromise fuel economy for other vehicle attributes that are their primary considerations, such as luxury performance and comfort, and the size of a vehicle. For truck buyers, valuations of fuel economy from pickup purchasers drag down consumer average WTP for a marginal increase of fuel economy to $233.

Considering long-term estimations, both car and light-truck buyers drastically undervalue fuel savings. According to the literature, if consumers substantially undervalue fuel economy, government interventions are necessary to help improve fuel economy (*30*). The results also suggest a substantially high discount rate for vehicle lifetime fuel savings, which is 44% by car consumers and 82% by truck buyers. These discount rates are higher than rates from previous studies (*9*). It was also found that at the discount rate of 7%, car consumers value discounted fuel savings for the first 3 years of ownership, whereas truck buyers value discounted fuel savings for the first year of ownership. These results are similar to Parry and Fischer's conclusion (*28*), which suggests a 3-year valuation of fuel economy. CAFE standards might be warranted to the extent that consumers value only short-term benefits from fuel economy improvement (*29*).

In a second-stage hedonic analysis, both education and age are positively significant in the estimations of results, indicating that vehicle consumers with higher education levels have a higher demand for fuel economy, and older buyers of cars have a higher fuel economy demand than younger car buyers. People with more education are more likely to be aware of the social effects of increased fuel consumption and GHG emissions. Therefore, it is reasonable for them to have a higher demand for fuel economy. Young buyers in Maine may prefer SUVs and large cars that have relatively lower fuel economy. Although they may not intentionally choose a lower fuel economy, their purchasing choice might compromise their demand for fuel economy. Income is not significant in both car and truck regression results. People with higher incomes prefer luxury features, and lower-income people might have more basic needs for the performance of vehicles.

Car consumers' net benefits from increasing current fuel economy from 25 to 35 mpg were calculated to be $2,232. Thus, higher CAFE standards are appropriate given the substantial net benefits from increasing fuel economy.

The major concern of increasing CAFE standards is the rebound effect. If more than one externality exists, then more than one policy instrument may be applied to adjust these externalities (*31*). The CAFE standards should be used with complementary policies. There are several suggestions for complementary policy instruments. For example, a mileage-based tax could be applied to mileage that exceeds a regulated level. A step tax that increases by mileage traveled could be the extension of this mileage-based tax. A registration fee based on mileage traveled may be another policy instrument. These policies might reduce the rebound effect in the form of offsetting the increased vehicle mileages traveled as a result of increased fuel economy. Congestion is another example of the rebound effect of CAFE standards. The alternative policy is varying tolls. If highway tolls are adjusted according to traffic peak periods and congested highway areas, people may choose to take an alternative route. These types of polices can help reduce congestion and thus reduce its related costs.

LIMITATIONS AND FUTURE DIRECTIONS

Implicit prices of fuel economy, vehicle weight, and power-to-weight ratio are endogenous because a certain level of marginal price of these attributes is associated with a certain level of attribute demand (*17*). In this case, SUR might produce inconsistent estimates of results. Future studies will examine the extent of inconsistency produced by the endogenous implicit prices in SUR. The proper identification of demand equations with endogeneity and the selection of valid instruments also will be further addressed in future studies.

Static analysis was used in analyzing fuel economy and introducing social benefits. It is assumed that consumers value fuel savings at the current fuel economy level. In practice, however, automotive fuel economy changes during a vehicle's lifetime, and fuel economy can be affected by the maintenance of automobiles. When social benefits are addressed, the change of externality costs and benefits across time during the life cycle of vehicles is ignored. These static assumptions might produce errors in the calculation of long-run fuel savings.

Safety indicators, which are primary considerations when vehicles are purchased, are not addressed in this study. This may create some inaccuracy in the estimation of results because of the omitted unobserved attributes.

Town-level income and education data are linked to a specific vehicle purchaser by town of residence, but income and education level vary among vehicle purchasers in town. This data limitation could lead to the insignificance of income.

Overall, this paper used an integrated two-stage hedonic model to estimate consumer valuation of fuel economy. In addition, this paper determined the cost-effectiveness of CAFE standards. CAFE standards are the most widely used performance-based regulations. Their cost-effectiveness, however, has not been well addressed in previous studies. Consumer valuations of fuel economy shed important light on further studies of policy design under different circumstances.

Of particular importance for this research, it was found that consumer valuations of fuel economy are not likely consistent with long-term benefits achieved from fuel economy improvement. This inconsistency could result from risk aversion and psychological heuristics and could lead to market failure. Future studies will further analyze what leads to behavior and market failures by using variables representing consumer preferences when they purchase vehicles. More conclusive suggestions could be brought to environmental and transportation arenas by comparing the social welfare from different policies. A combination of policies that help reduce the rebound effect from CAFE standards should be further analyzed. More-detailed and comprehensive data sets might allow estimation of consumer valuations of fuel economy across time and space. The valuations of fuel economy from vehicle purchasers in California, where air pollution is more serious, may be different than those in Maine.

ACKNOWLEDGMENTS

The authors thank Gary Hunt and Hsiang-tai Cheng for their assistance in econometric modeling. The authors also thank David Greene for his comment on discount rates. The authors appreciate the constructive comments of the anonymous reviewers.

REFERENCES

1. *Energy Independence and Security Act of 2007.* http://thomas.loc.gov/cgi-bin/bdquery/D?d110:1:./temp/~bdcDQ6::|/bss/110search.html|. Accessed Feb. 5, 2008.
2. *Average Fuel Economy Standards, Passenger Cars and Light Trucks, Model Years 2011–2015.* NHTSA, 2007.
3. Increased Corporate Average Fuel Economy Standards. *Energy Independence and Security Act of 2007.* http://www.govtrack.us/congress/bill.xpd?bill=h110-6. Accessed Aug. 1, 2008.
4. Griliches, Z. Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In *The Price Statistics of the Federal Government,* National Bureau of Economic Research, New York, 1961.
5. Goodman, A. C. Willingness to Pay for Car Efficiency: A Hedonic Price Approach. *Journal of Transport Economics and Policy,* Vol. 17, No. 3, 1983, pp. 247–266.
6. Berry, S., S. Kortum, and A. Pakes. Environmental Change and Hedonic Cost Functions for Automobiles. *Proceedings of the National Academy of Sciences Online,* Vol. 93, No. 23, 1996, pp. 12731–12738.
7. Espey, M., and S. Nair. Automobile Fuel Efficiency: What Is It Worth? *Contemporary Economic Policy,* Vol. 23, No. 3, 2005, pp. 317–323.
8. Bhat, C. Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research Part B,* Vol. 35, No. 7, 2001, pp. 677–693.
9. McManus, W. *Can Proactive Fuel Economy Strategies Help Automakers Mitigate Fuel-Price Risks?* UMTRI Automotive Analysis Division, University of Michigan, Ann Arbor, 2006.
10. Train, K., and C. Winston. Vehicle Choice Behavior and the Declining Market Share of U.S. Automakers. *International Economic Review,* Vol. 48, No. 4, 2007, pp. 1469–1496.
11. Baranzini, A., J. Ramirez, C. Schaerer, and P. Thalmann. *Hedonic Methods in Housing Markets.* Springer, New York, 2008.
12. Court, A. T. *Hedonic Price Indexes with Automotive Examples in Dynamics of Automobile Demand.* General Motors, New York, 1939.
13. Pickering, J., J. A. Harrison, B. C. Isherwood, J. J. Hebden, and D. H. Hitchen. Are Goods Goods? Some Empirical Evidence. *Applied Economics,* Vol. 5, 1973, pp. 1–18.
14. Hogarty, T. F. Price-Quality Relations for Automobiles: A New Approach. *Applied Economics* Vol. 7, 1975, pp. 41–51.
15. Triplett, J. E. Automobiles and Hedonic Quality Measurement. *Journal of Political Economy,* Vol. 77, No. 3, 1969, pp. 408–417.
16. Ohta, M., and Z. Griliches. Automobile Prices and Quality: Did the Gasoline Price Increases Change Consumer Tastes in the US? *Journal of Business and Economic Statistics,* Vol. 4, No. 2, 1986, pp. 187–198.
17. Champ, P. A., K. J. Boyle, and T. C. Brown. A Primer on Nonmarket Valuation. Kluwer Academic Publishers, Dordrecht, Netherlands, 2003.
18. Brown, J. N., and S. R. Harvey. On the Estimation of Structural Hedonic Price Models. *Econometrica,* Vol. 50, No. 3, May 1982, pp. 765–768.
19. Epple, D. Hedonic Prices and Implicit Markets: Estimating Demand and Supply Functions for Differentiated Products. *Journal of Political Economy.* Vol. 95, No. 1, 1987, pp. 59–80.
20. Beron, K., J. Murdoch, and M. Thayer. The Benefits of Visibility Improvement: New Evidence from the Los Angeles Metropolitan Area. *Journal of Real Estate Finance and Economics,* Vol. 22, No. 2/3, 2001, pp. 319–337.
21. Rosen, S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy,* Vol. 82, No. 1, 1974, pp. 34–55.
22. Key Automotive Data. Ward's Automotive Group. http://wardsauto.com/keydata/. Accessed July 1, 2008.
23. Vehicle Registration and Titles. Bureau of Motor Vehicles. http://www.maine.gov/sos/bmv/registration/. Accessed Aug. 1, 2007.
24. Information Resource of Maine. http://www.maine.gov/informe/. Accessed Aug. 1, 2007.
25. *International Energy Outlook 2007.* Energy Information Administration. http://www.eia.doe.gov/oiaf/archive/ieo07/index.html. 2007. Accessed Feb. 4, 2008.
26. *Corporate Average Fuel Economy for MY 2011 Passenger Cars and Light Trucks. Final Regulatory Impact Analysis.* NHTSA. http://www.nhtsa.dot.gov/portal/site/nhtsa/menuitem.43ac99aefa80569eea57529cdba046a0/. Accessed Feb. 2, 2009.
27. Dreyfus, M., and K. Viscusi. Rates of Time Preference and Consumer Valuations of Automobile Safety and Fuel Efficiency. *Journal of Law and Economics,* Vol. 38, No. 1, 1995, pp. 79–105.
28. Parry, I. W. H., and C. Fischer. *Do Market Failures Justify Tightening Corporate Average Fuel Economy (CAFE) Standards?* Resources for the Future, Washington, D.C., 2005.
29. Fischer, C., W. Harrington, and I. W. H. Parry. *Economic Impacts of Tightening the Corporate Average Fuel Economy (CAFE) Standards.* Resources for the Future, Washington, D.C., 2004.
30. Greene, D., P. Patterson, M. Singh, and J. Li. Feebates, Rebates and Gassguzzler Taxes: A Study of Incentives for Increased Fuel Economy. *Energy Policy,* Vol. 33, No. 6, 2005. pp. 757–775.
31. Rubin, J., P. N. Leiby, and D. L. Greene. Tradable Fuel Economy Credits: Competition and Oligopoly. *Journal of Environmental Economics and Management,* Vol. 58, No. 3, 2009, pp. 315–328.

The Travel Analysis Methods Section peer-reviewed this paper.

Time to Play?

Activity Engagement in Multiplayer Online Role-Playing Games

H. S. Mahmassani, R. B. Chen, Y. Huang, D. Williams, and N. Contractor

The online session behavior of users in a multiplayer online role playing game is examined. Three dimensions of online sessions are studied: frequency of sessions, distribution of start and stop times, and duration. These measures are modeled and examined econometrically to understand their relationships with attributes of the users and their avatars. Frequency is analyzed by using a negative binomial regression to uncover relationships with contributing factors, such as the experience level of players and the number of avatars per user account. Start and stop time distributions are examined aggregately with discrete choice models that capture the probability of starting a session in relation to the sociodemographics of users. Duration models are estimated by using data on the characteristics of events that occur within the duration and other user sociodemographics and player attributes. Since sessions are not explicitly indicated within the data set, sessions were stitched together on the basis of an assumed threshold for offline duration. Overall, these models and results provide insight into temporal characteristics of online sessions, including their aggregate traffic characteristics, such as frequency, headways, and duration length. These results also provide information on the nature of online game playing from a user perspective.

The Internet has had profound effects on many aspects of life and has transformed, in complex ways yet to be understood, the manner in which individuals go about activities normally pursued in the physical world, including work, shopping, maintenance, and leisure. These transformative effects have been more evident in some segments of the population than others, for example, youth for whom virtual social networking appears to be an essential component of overall social engagement, strongly intertwined with its physical counterpart. Although analysts have shed some light on the role of information and communication technologies for work-related activities, from both a work management perspective (*1–4*) and a travel or spatial location perspective (*5–7*), much less is known about the extent to which the virtual world has become an integral component of individuals' leisure time.

Game playing has been an important use of computers since the introduction of personal computers in the early 1980s. However, online multiplayer games gained popularity only with the advent of broadband Internet in homes. Although online gaming in its early forms (e.g., Oregon Trail) gained favor with hardcore computer users, involvement was limited to those with access to networking capabilities in the academic or software development environments. Near-ubiquitous residential broadband access has been accompanied by increasing engagement in online multiplayer games over the Internet, initially by the population segment most commonly associated with leisure seeking and technology adoption, that is, young males, and now by growing numbers of people from a much broader demographic base (*8, 9*). In addition to their evident gaming features (e.g., scoring points by winning or accomplishing certain milestones), online games have a distinct social interaction component that may be just as significant a motivation for participating in such activity. As such, the lines between social networking and online gaming have become more blurred, with social networking sites offering many gaming options (e.g., Mafia Wars on Facebook, the quintessential social networking site) and online gaming environments acquiring an identity and presence of their own for virtual socialization and even business transactions (e.g., Second Life).

Because so much time is spent by so many people in online gaming environments, and because of its particular significance to younger cohorts, characterizing the patterns of online gaming as an important activity in individuals' daily and weekly activity engagement patterns is a logical pursuit. The field of travel behavior has long recognized the integral connections among time use, activity engagement, and travel. As such, characterizing online gaming engagement will contribute to understanding how information and communication technologies are transforming individual time use and activity patterns in critical ways and contribute to improved analysis tools for transportation policy that recognize the importance of the virtual world in determining activity and travel behavior in the physical world.

A major obstacle to understanding engagement in online gaming activity is the difficulty of obtaining data on participation in these activities. Such information is not typically obtained in traditional travel and activity survey diaries, and reliance on self-reported information on this matter is unsupported—anecdotal evidence suggests that gamers are apt to either misperceive or underreport the amount of time spent in an online gaming environment (*10–12*). However, that these games are computer based means that an exact log of events and activities conducted by a player in the virtual gaming environment could be recorded directly. Such information would therefore be available to the entity (server) hosting the game, although it would of course be subject to strict privacy rules. The work presented in this paper is based on such data, obtained for a game called Everquest II, which is a popular multiuser online gaming environment with adherents across the world. A sample of event logs for this game formed the basis of the analysis of activity engagement conducted in this study.

H. S. Mahmassani and R. B. Chen, Transportation Center, Northwestern University, 600 Foster Street, Evanston, IL 60208-4055. Y. Huang and N. Contractor, Science of Networks in Communities (SONIC) Lab, Northwestern University, 2-133 Frances Searle Building, 2240 Campus Drive, Evanston, IL 60208-2952. D. Williams, University of Southern California, Suite 121E, 3502 Watt Way, Los Angeles, CA 90089-0281. Corresponding author: H. S. Mahmassani, masmah@northwestern.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 129–137.
DOI: 10.3141/2157-16

The primary aspects of online gaming engagement addressed in this study consist of the frequency, timing, and duration of virtual gaming spells, viewed over a weekly horizon. These provide a ready picture of the extent and significance of online gaming relative to an individual player's total activity pattern, including both weekday and weekend engagement. Additional insight into factors that determine such engagement, both intrinsic to the game itself as well as extrinsic attributes of the players, particularly age and gender, are provided. A foundation is set for examining the dynamics of game activity engagement by studying the extent to which individual engagement increases or decreases over time and the factors that help explain such engagement. To the authors' knowledge, this is the first analysis conducted from an activity participation perspective for any of the major online game environments.

DATA DESCRIPTION AND CONCEPTS

The data set consists of event logs for all players subscribing to the Everquest II game across four servers in 2006. Servers are located in Southern California and thus all time stamps are Pacific Standard Time. Event logs consists of the time stamp, location within the virtual world, and relevant characteristics of different event types within the game, ranging from economy-related activities, such as purchasing items, to death. For the purpose of this analysis, a sample of 750 individuals was drawn from the population of game players for the first week of August 2006. However, because of the need for the players' geographic locations to adjust for time zone differences, only 552 individuals were suitable for analysis. For each player, the time stamps of 12 of the most frequent types of events were analyzed to determine the online sessions of activity engagement of players. One activity carried out in the virtual environment does not necessarily correspond to one event in the log but possibly several. For example, the slaying of a monster may lead to an increase in experience points and at the same time may move the player to the next stage of a quest or to multistage tasks that require a sequence of events to be completed.

Since the start and finish of sessions or gaming spells are not directly recorded in the log, these are determined on the basis of an assumed threshold for offline periods, in which the duration of inactivity that exceeds this threshold indicates the player is offline. For this study, this threshold was assumed to be 20 min. Given the chronological ordering of events according to their time stamps, the time elapsed between events was determined for all events that occurred over the first week of August and compared against the 20-min threshold. Gaps that exceeded the assumed offline threshold (20 min) were considered offline durations, and the durations between offline durations were considered online session events, with a start time and duration. This method of reconstructing online sessions is called session stitching and has been used in other fields that examine the online sessions of players (*8, 13*). These studies used a similar threshold value. A look at the descriptive statistics of the player sample reveals some insight into the age range and sociodemographics of the players, as well as information regarding their avatars and accounts. The distribution of players according to gender and residential location is shown in Figure 1. Although, increasingly, a much broader demographic base engages in these games, players are predominantly male, at 82% of the sample, indicating that these games, or at least multiplayer online role-playing games, have reached a wider demographic but not yet to a significant level. Furthermore, 84% of the players in the sample reside in the United States. The location of players' residences has implications for the coordination of games among groups of players. If players are distributed globally, then coordinating efforts in the game, such as slaying major monsters, may be problematic.

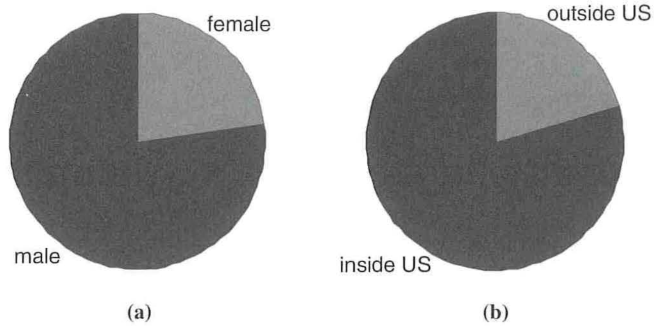


FIGURE 1 Distribution of (*a*) gender and (*b*) location of players.

Figure 2 shows the distribution of players for age, age of avatars, age of game accounts, and number of avatars each player has. The age range of players is 15 to 71 years. The figure shows that most of the players are in their 20s and 30s, suggesting that relatively few are teenagers or children who still attend school. This suggests that players may have rigid schedules; if jobs occupy most of the day, free time for these game activities would be found either after or before work. The age of the many of the avatars in the sample does not exceed 1 year. Compared to the age of the players' game accounts, which has a median value of about 2.1 years, the age of avatars is shorter overall. This suggests that many players associate multiple avatars with their accounts over time. This is important to the investigation of learning behaviors in the game. Although the observed avatars may have existed for a relatively short time, the player's account may span several years, suggesting these accumulated experiences would be reflected in the avatar. The number of avatars a player has at any one time ranges from one to 10 with one being the mode.

Overall these statistics are consistent with those from other online games. Two counterintuitive trends observed in this game and other games concern gender and mean age of players. Other studies have found that although females make up a minority of players, they are more intensive in their game playing relative to their male counterparts (*8*). Also, the mean age is counterintuitively around 35 years. One would expect that players for an online game would be younger. However, this further suggests that engagement in these games is not limited to the youth segment of the population (*9*).

EXPLORATORY ANALYSIS

This section presents an exploratory analysis of online gaming activity engagement and shows basic statistical summaries and tabulations of the key aspects of interest to the study. A total of 5,957 sessions were observed after the session-stitching process, described previously, was completed. This number reflects the exclusion of sessions for which the location of the player was unavailable and thus the start and stop times could not be adjusted for differences in time zone. The start and finish times of online sessions provide a measure of the number of sessions that start that could vary with the characteristics of the player and of the session itself. For example, individuals who work during the day are less likely to engage in online games for

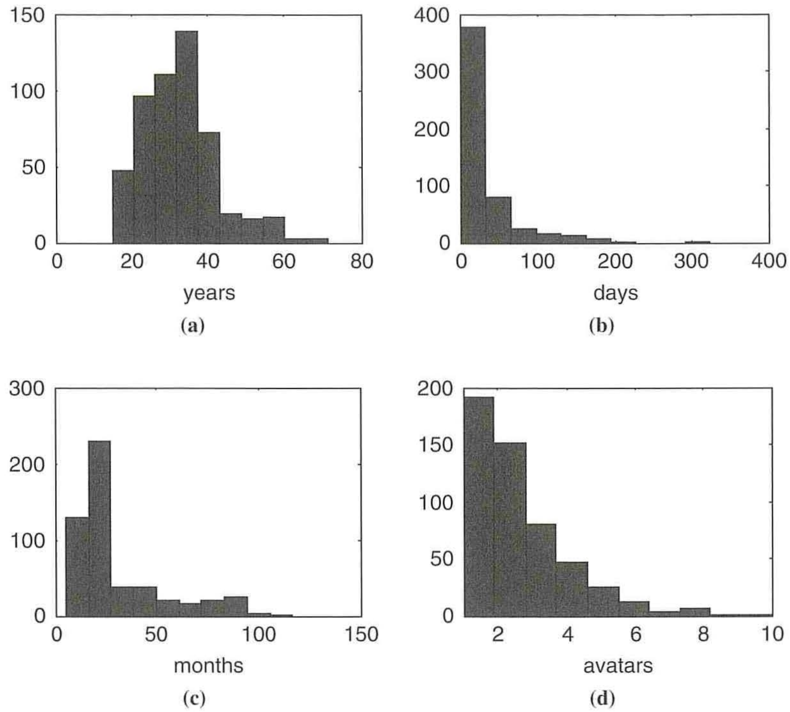


FIGURE 2 Distribution of players by (*a*) age of players (years), (*b*) age of avatar (days), (*c*) age of account (months), and (*d*) number of avatars.

long durations. The number of sessions at a given time is important for systems management to ensure sufficient resources and capacity. For the users or players, information on the session start times of other players may be useful for activities that require networking or making social connections with other players. For online multiplayer role-playing games, some activities require the efforts of multiple players, and thus novice players who do not know many or any other players may try to be online to increase the likelihood of connecting with others. For travel behavior analysts, there is a need to understand and better position online game-playing activities with others, especially those involving multiple players or social networks.

The first characteristic examined is the cumulative number of users online in the virtual world environment over time. This provides a measure of the number of users playing the game over time and the number of sessions that begin relative to the number finishing. Figure 3 shows the cumulative number of session starts and stops for 1 week. A noticeable trend is that across all days, the total number of sessions starts out low in the first half of the day (12:00 midnight to 12:00 noon) and then rises sharply and decreases sharply after midnight. For Friday, Saturday, and Sunday this sharp rise occurs slightly earlier relative to the other days of the week. This is indicated in Figure 3*b*, where the number of sessions raises and falls cyclically over a week. Given that individuals typically sleep during first quarter of a day (12:00 midnight to 6:00 a.m.), this low period is reasonable. The shift in the sharp rise in sessions on the weekend could be explained by more free time for playing during the earlier part of the day, relative to weekdays. However, without information on the other activities in which players engage during the day, this cannot be known. This is indicated by the wider peaks on Friday, Saturday, and Sunday relative to other days, as indicated in Figure 3*b*.

To characterize the choice of session start time, the percentage of players in the sample with start times falling within specific hours of the day are shown in Figure 4. The distribution of sessions start times reveals many characteristics of the relationship between the time of day and the decisions of users to start a session.

First, for all days of the week, there are more session starts in the later part of the day relative to the early part. This is also shown in Figure 3. A comparison of Friday to the other days shows a noticeable spike in the number of session starts in the middle of the day. A comparison of the weekend with weekdays shows more session starts between 10:00 a.m. and 3:00 p.m. on weekends relative to weekdays. These differences indicate possible differences for day of the week. More session starts in the middle part of weekend days may indicate that players have other mandatory activities during that time on weekdays, such as work, and thus show fewer observations between 10:00 a.m. and 3:00 p.m. The strong peak on Friday may indicate that players prefer to start playing during the afternoon relative to other periods of the day. Second, across all days, a midday peak occurs after 10:00 a.m. For Fridays, this is the top peak. Possible reasons range from playing during a break from mandatory activities, such as lunch breaks, to coordination with other players. However, without a closer examination of the characteristics of these users and of the sessions themselves, the motivations behind these session starts cannot be known.

The final characteristic examined in this exploratory analysis was the distribution of session durations across different hours of the day. Figure 5 shows the distribution of mean duration of the sessions starting in the corresponding times of day. First, the results indicate that the shortest durations occur in the morning period just before 10:00 a.m., except on weekends, when the shortest durations are sessions that start just after midnight. Second, the durations of

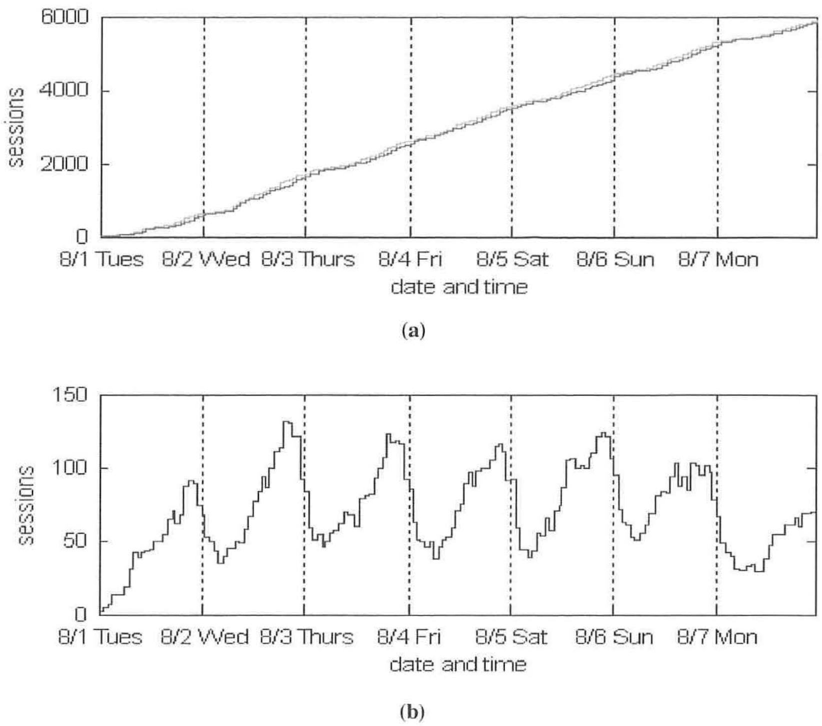


FIGURE 3 (*a*) Cumulative session starts and stops over 1 week and (*b*) number of sessions over time.

sessions started during Friday and the weekends are higher than those during the weekdays. The durations reach a peak of about 100 min on weekdays and about 150 min for other days. This is possibly because there is more free time to play these games on nonweekdays. A second curious characteristic of the mean session durations is that for Friday and weekdays, the sessions that start between 5:00 and 10:00 a.m. are the shortest, but on the weekends these sessions have some of the longest durations. A possible explanation for this is that players who start during these periods are pursuing activities that require more time, or perhaps these players enjoy playing the game for longer durations relative to players that start at other times.

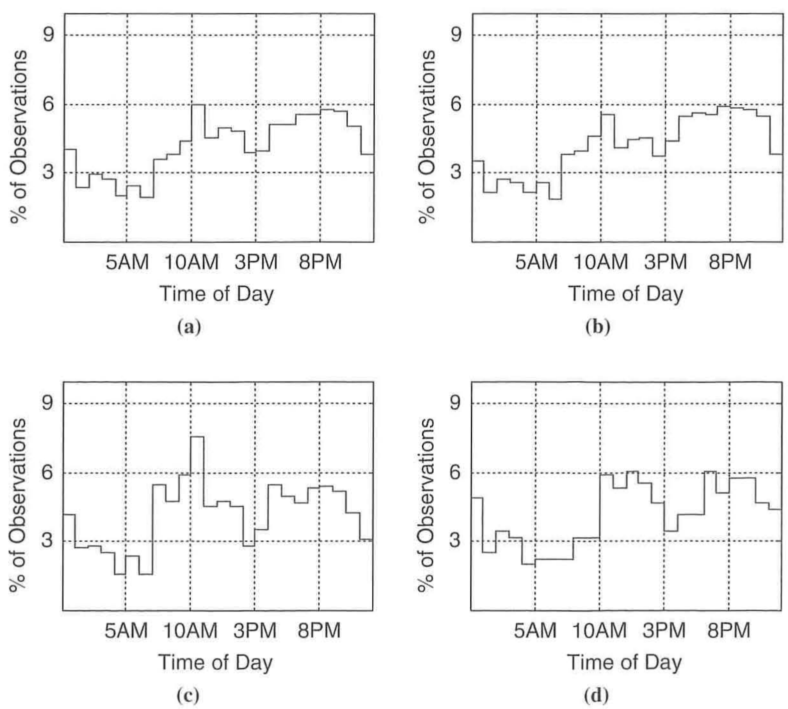


FIGURE 4 Distribution of session start times (*a*) all days, (*b*) weekdays, (*c*) Fridays, and (*d*) weekends.

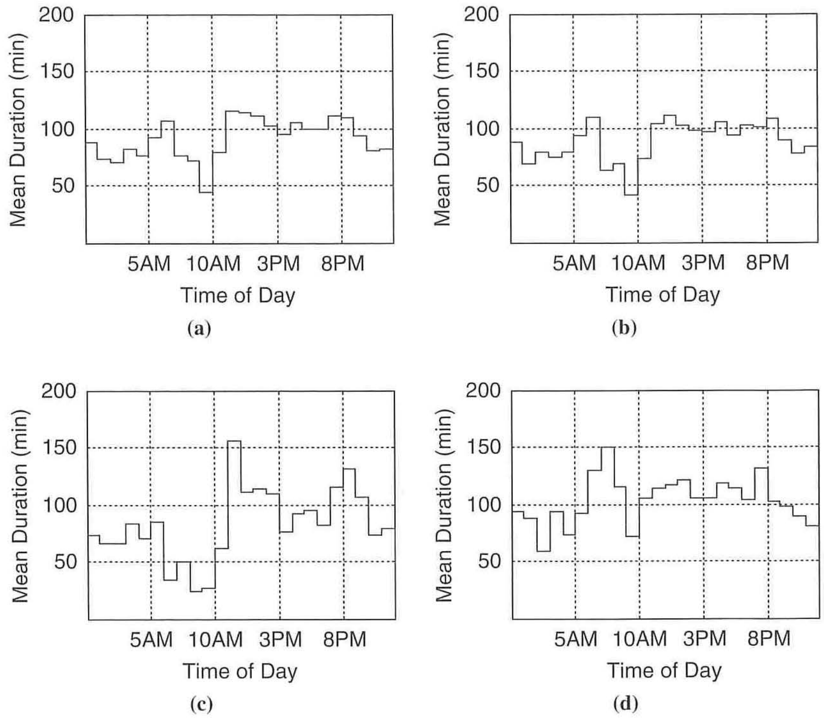


FIGURE 5 Distribution of mean session durations (*a*) all days, (*b*) weekdays, (*c*) Fridays, and (*d*) weekends.

Overall, the distribution of session start and stop times and the mean durations of sessions indicate that the level of players in the environment varies over time. This level is cyclical with low levels during the early parts of the day and rising levels later in the day. This appears to correspond well for common work and sleep schedules, in which players work during the day and sleep at the end of the day. For Fridays and weekends, this peak starts earlier relative to weekdays and lasts longer. In general, there are more session finishes in the early part of a day (after midnight) than in other parts.

A closer examination of the characteristics of users and sessions may reveal more insights about the players and their engagement frequencies, choice of start time, and durations. The next three sections examine the relationships among these characteristics of sessions, and the attributes of the players and their avatars or characters.

FREQUENCY OF ONLINE ACTIVITY ENGAGEMENT

The frequency of activity engagement is further examined by developing and estimating a negative binomial model that relates frequency to various attributes of the player, including experience level, sociodemographic characteristics, and attributes of the avatar, such as age. Frequency was measured in sessions per week. A negative binomial model was chosen to account for possible overdispersion where the mean is less than the variance (*14*). Overdispersion occurs for a variety of reasons related to the process generating the count data. The most common reason in many studies is that variables influencing the Poisson rate have been omitted. The estimation results for this model are shown in Table 1. All 688 sampled individuals were used in the estimation, since frequency does not require information on the time zones of the individuals. First, the overdispersion parameter is significant, indicating that the variance is larger than the mean. Second, the restricted log likelihood test suggests that the fitted model is better than the model with just the constant term. The McFadden pseudo R-squared value is 0.33, suggesting this model may not fit the data very well.

Overall the estimates indicate that the sociodemographics of the player have little significance with the frequency of engagement relative to attributes of the avatar and account. This is indicated by the low t-statistics for the indicator for residing in the United States, resulting in its being dropped. Playing frequency may have no relation to cultural differences. In general, as the age of the player increases, the propensity to engage in more-frequent playing decreases, although this parameter is not significant. As the age of the avatar increases, the propensity to more frequently engage in sessions increases, then decreases. There are several plausible explanations for this, such as the design of the game. In Everquest II, undoubtedly the tasks may become more involved, requiring more time, which may require all the budgeted game time from players. Players with young avatars may be faced with tasks that are easy and require less duration. Thus, these players can engage in the game more frequently

TABLE 1 Negative Binomial Model of Online Activity Engagement

Variable	Estimate	t-Statistic
Constant	2.4401	27.581
Age player (years)	−0.0051	−2.802
Age avatar (days)	0.0100	6.521
Age avatar[a] (115 days < age < 231 days)	−0.0077	−4.743
Age avatar[a] (age >231 days)	−0.0001	−2.455
Age account (days)	−0.0001	−1.906
Overdispersion parameter	0.4868	13.719

NOTE: Number of observations = 628; log likelihood (constants) = −2,147.258; log likelihood (convergence) = −2,115.803; and chi-squared = 2,127.608.
[a]Binary (1/0) variable.

within the budgeted game time. Alternatively, players with very skilled avatars may lose interest in the game and thus play less frequently. Both explanations are plausible but require further investigation into the dynamics of the player and avatar over time, to possibly distinguish how the two characteristics (player and avatar) interact and affect the frequency of engagement. As the age of a player's account increases, propensity to engage decreases. Similar to age of the avatar, a possible explanation is lack of interest from the player. However, closer examination of the motivations behind player perception of the game is required to fully address this question.

TIMING OF ACTIVITY SPELLS

The timing and duration of activity spells is then analyzed through specification and calibration of a "time of day for play" choice model, as well as a hazard-based model relating the duration of play sessions to various attributes of the player and his or her experience in this environment. A multinomial logit model form was used to estimate this choice. In the sample, the number of observations per person varied from one to several, if the person played the game several times a week. Not accounting for possible serial correlation may have implications on the coefficient estimates. Each player had a choice among eight periods for starting a session, the first period starting at 12:00 midnight and the last starting at 9:00 p.m. These results are shown in Table 2.

Period 8 (10:00 p.m. to 12:00 midnight) was used as the base or reference alternative. The results from the choice model reveal some important findings about the characteristics of the player and their choice of online session start time. First, with respect to age, only between 12:00 midnight and 12:00 noon do differences arise. Between 12:00 midnight and 6:00 a.m. there is a negative propensity toward playing the game for those age 36 years and older and between 3:00 a.m. and 6:00 a.m. for ages 25 to 35. This is shown in Tables 2 and 3; Table 3 shows the coefficient values across attributes across periods. This negative propensity is plausible given that most players would be sleeping between 12:00 midnight and 6:00 a.m. Also, the results show that this negative propensity starts earlier (12:00 midnight to 3:00 a.m.) for ages 36 and older, relative to other ages. Between 9:00 a.m. and 12:00 noon, players ages 36 and older are more likely to play relative to other ages, and between 6:00 and 9:00 a.m. ages 26 and older are more likely. Although normally players may have mandatory activities during these periods, such as work, players may be logging on during a lunch break. From Figure 4, it appears a large number of starts occur around 11:00 a.m. The results show that males are more likely to play during the late-night hours (12:00 midnight to 6:00 a.m.) compared to females. Estimates show that players residing in the United States have a strong propensity for playing in the early morning and a negative propensity for playing between 3:00 and 6:00 p.m. Overall, individual characteristics appear to matter less for periods after 12:00 noon for playing the game.

Fridays appear to positively affect playing between 12:00 midnight and 3:00 a.m. and between 6:00 a.m. and 12:00 noon. This is reasonable considering that many individuals do not need to work the next day, Saturday. The weekend positively affects playing throughout the day except the period between 3:00 and 6:00 p.m., which shows a negative sign. One reason is that individuals would rather spend their afternoons, whether on the weekend or a weekday, doing something besides playing the game. Also, nighttime in general appears to be a consistent period for playing regardless of day of the week.

TABLE 2 Choice of Start Time

Variable	Estimate	*t*-Statistic
Alternative Specific Constants		
Period 1	−0.6645	−4.613
Period 2	−1.1283	−6.402
Period 3	−1.3515	−8.942
Period 4	−0.0650	−1.162
Period 5	−0.1330	−2.360
Period 6	0.3096	3.382
Period 7	0.0314	0.414
Period 2 Variables	3:00 a.m.–5:00 a.m.	
Age 26–35 years (1/0)	−0.2154	−1.987
Age 36 years and older (1/0)	−0.3074	−3.294
Male (1/0)	0.1457	1.995
Reside in United States (1/0)	0.5137	4.684
Age avatar (days)	0.0352	7.395
Age avatar >30 days	−0.0234	−6.154
Age avatar >60 days	−0.0084	−3.694
Age account (days)	−0.0009	−4.648
Age account >2 years	0.0004	4.383
Period 4 Variables	9:00 a.m.–11:00 a.m.	
Age 36 years and older (1/0)	0.1464	1.826
Friday (1/0)	0.3370	4.220

Variable	Estimate	*t*-Statistic
Period 1 Variables	12:00 midnight–2:00 a.m.	
Age 36 years and older (1/0)	−0.3074	−3.294
Male (1/0)	0.1457	1.995
Weekend (1/0)	0.2640	2.866
Friday (1/0)	0.2640	2.866
Age avatar (days)	0.0352	7.395
Age avatar >30 days	−0.0234	−6.154
Age avatar >60 days	−0.0084	−3.694
Age account (days)	−0.0009	−4.648
Age account >2 years (days)	0.0004	4.383
Period 3 Variables	6:00 a.m.–8:00 a.m.	
Age 26 years and older	0.6050	5.037
Weekend (1/0)	−0.2391	−2.201
Friday (1/0)	0.3370	4.220
Reside in United States (1/0)	0.5137	4.684
Period 5 Variables	12:00 noon–2:00 p.m.	
Weekend (1/0)	0.2293	2.755
Period 6 Variables	3:00 p.m.–5:00 p.m.	
Weekend (1/0)	−0.2846	−3.171
Reside in United States (1/0)	−0.3240	−3.461
Period 7 Variables	6:00 p.m.–8:00 p.m.	
Male (1/0)	0.1457	1.995

NOTE: Number of observations = 524; log likelihood (constants) = −12,179.72; log likelihood (convergence) = −11,973.45; and rho-squared = 0.00989.

TABLE 3 Coefficient Values for Attributes Across Periods

Period	12:00 midnight–2:00 a.m. 1	3:00 a.m.–5:00 a.m. 2	6:00 a.m.–8:00 a.m. 3	9:00 a.m.–11:00 a.m. 4	12:00 noon–2:00 p.m. 5	3:00 p.m.–5:00 p.m. 6	6:00 p.m.–8:00 p.m. 7
Age 26–35 years		−0.2154	0.6050				
Age 36 years and older		−0.307		0.1464			
Male		0.1457					0.1457
Friday	0.2640			0.3370			
Weekend					0.2293	−0.2846	
Reside in United States			0.5137			−0.3240	

Overall, in relation to time of day for playing, nighttime and afternoon appear to be periods that are not strongly influenced by characteristics such as age. Only for the late-night period do players of different ages exhibit different propensities toward playing. This is reasonable considering that nighttime is typically free from work across all age groups and that more individuals do not sleep before 9:00 p.m. However, for the late night, depending on activities for the following day and on physiological reasons, players may vary in their playing. Some individuals need to sleep earlier relative to others. The results suggest that players older than 36 are less likely to stay up and play relative to players who are 26 to 35 years old. Fridays appear to show higher propensities only between 12:00 midnight and 12:00 noon, likely because players do not need to work the next day.

Another set of dimensions that provide insight into the time of day for playing is the age of the account and the avatar. These factors appear to influence choice only between 12:00 midnight and 6:00 a.m. The relationship between utility and age of the avatar and the account between 12:00 midnight and 6:00 a.m. is shown in Figure 6. The figure shows that the propensity for playing between 12:00 midnight and 6:00 a.m. increases with age of the avatar, but this rate decreases over time. Assuming that age is a measure of experience gained by the avatar, the results suggest that as players gain experience, they are more likely to play during late hours. This suggests that players with avatars with high experience levels tend to play late at night, possibly to coordinate with other players more easily. The age of the account, unlike the age of the avatar, is a measure of how long a player has been playing the game. A player may have an avatar with a low age but have an account that is many days old with multiple prior avatars. Figure 6 shows that the older an account, the less likely the player will play during late hours. A possible reason for this is loss of interest in the game over time. With lower interest, playing the game may not take much precedence during the late hours over other activities. This is similar to the negative coefficient for age of the account in the frequency model. Thus, over time players may lose interest and play the game less.

DURATION OF GAME ACTIVITY SPELLS

A duration model of game activity spells was estimated to show the relationship with characteristics of the player. These estimation results are shown in Table 4.

Two different parametric models were estimated to account for duration effects. The exponential model assumes a constant hazard function, and thus the probability of a gaming activity spell ending is independent of the time spent playing. For an exponential model, the distribution parameter in the hazard (P) is equal to 1.0. The Weibull model shows P to be very different from 1.0 statistically, so the exponential model is not valid. The positive value on P for the Weibull distribution indicates the hazard increases with increasing duration, meaning the player is more likely to exit the gaming activity spell as the duration increases.

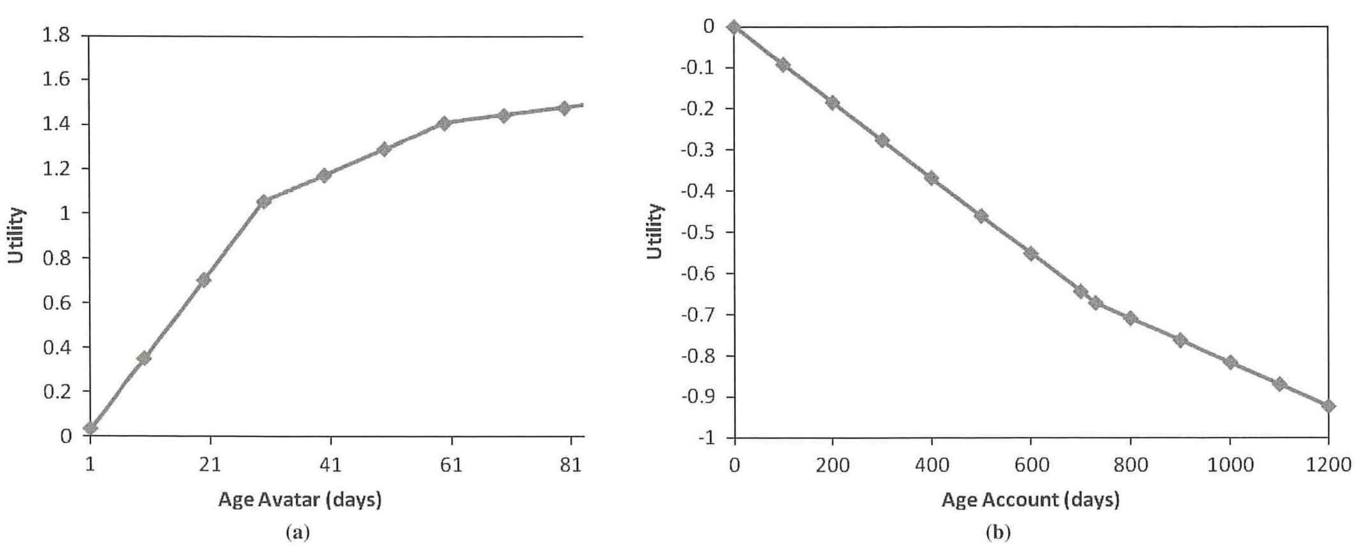


FIGURE 6 Relationship between utility and (*a*) age of avatar and (*b*) age of account, 12:00 midnight to 6:00 a.m.

TABLE 4 Hazard Function Parameter Estimates for Duration of Game Activity Spells

Parameter	Exponential		Weibull	
	Estimate	*t*-Statistic	Estimate	*t*-Statistic
Constant	8.5446	228.420	8.3096	111.009
FRIDAY (1/0)	−0.0559	−2.102	−0.1162	−2.194
WEEKEND (1/0)	0.1612	6.511	0.1804	3.625
LOC_US (1/0)	−0.1598	−5.421	−0.1701	−2.932
MALE (1/0)	0.1994	7.906	0.2186	4.278
Age 26–35 years (1/0)	0.0544	2.612	0.0652	1.559
Age account >2 years (1/0)	−0.1833	−8.122	−0.1944	−4.300
Age account >5 years (1/0)	−0.1411	−4.253	−0.1417	−2.138
Age avatar >1 month (1/0)	0.3007	11.658	0.3057	5.841
Age avatar >2 months (1/0)	−0.2237	−6.862	−0.2858	−4.418
P (distribution parameter)	1.0000	—	0.6824	84.975
Lambda	0.0002	18.000	0.0023	23.000
Log likelihood at convergence	−12,470.84		−11,674.69	
Log likelihood (constants)	−12,606.49		−11,749.39	
Number of observations (persons)	524		524	

The model estimates indicate that durations tend to be longer on Fridays and shorter on Saturday relative to other days. Given the results in Tables 2 and 3, this difference may be explained by considering the relationship between gaming activities with other activities. Tables 2 and 3 show a negative coefficient for playing during the afternoon on weekends. A possible explanation is that playing these games does not take precedence over other activities. Notice also that much of the game playing occurs at night, further suggesting that these games typically do not take a mandatory or priority role. The parameters for the duration model suggest that on the weekends durations are smaller, supporting the reasoning that given other activities and the time to pursue them, these gaming activity spells are adjusted for time of day and duration. However, further information on other activities the individual faces is needed to better understand this relationship. The estimation results also show that males tend to play shorter durations than do females. This is in line with other studies on the same data set that show that females who play these games tend to be more intense than their male counterparts, although statistically they are a minority (*8*).

Players with old accounts tend to play longer relative to players who just opened their accounts. Also, younger (less-experienced) avatars tend to play for shorter durations relative to more-experienced avatars. This may be because less-experienced characters or avatars have less-complicated tasks relative to more-experienced avatars that are further along in the game storyline. Individuals between 26 and 35 years old tend to play for shorter durations. Without knowledge of other activities in the players' schedules, the relationship of these durations to time budgeted for other activities cannot be fully investigated and determined.

CONCLUDING COMMENTS

Increasingly, members of all sociodemographic groups engage in online multiplayer games. This engagement is further encouraged by the popularity of online social networking, which has become ubiquitous in the daily activities of youth and adults. This study examined the behaviors of players in an online multiplayer game called Everquest II. The results of this study apply to this specific game and may not be generalized to all online multiplayer role-playing games. The results suggest that playing occurs mostly at night, and for late-night periods, the start of sessions varies with individual characteristics such as age. The results also show that day of the week plays a role, although without more information on other activities the reasons for these differences cannot be determined. The results show that engagement in these games is related to the experience of players. For example, players whose characters have more experience tend to play longer durations. Also players with very old accounts tend to play less frequently, possibly because of declining interest in the game.

In general, results suggest that engagement in these gaming activities is related to three factors: (*a*) attributes of the player, (*b*) attributes of the game, and (*c*) other activities in a person's activity schedule. The second factor adds a layer of complexity to study of these gaming activities. Players represent themselves in the game not directly but rather through an avatar or character that exists in the virtual environment. Thus to fully understand the behaviors of these players for the dimensions investigated in this study, a better representation is needed of the relationships among the player, his or her avatar, and the other activities of the players.

Studying the gaming activities of individuals is important to travel behavior, especially in the context of telecommunications. These games may be complementary to other online activities such that they trigger other activities that require trips, such as shopping or socializing with friends. An important attribute of these online games is that there is a social networking component that plays a vital role in the players' gaming experience. For example, some tasks, such as fighting monsters, require a group of players to be successful. Social networking components include guilds, which are similar to social clubs in the virtual environment. Thus, these online role-playing games are quickly becoming potential substitutes for socializing with other friends.

Although the data do not allow a complete profiling of who plays and how much, they contribute to an understanding of the significance of these activities to a broad cross section of the population that transcends popular stereotypes. Of course, one would also like to know a lot more about why people play these games and the different dimensions (recreation, social, skill, accomplishment, etc.) that such engagement fulfills for the individual players. In particular, the connection between activities performed virtually and those conducted in the physical world are of importance to researchers into travel behavior. Definitive conclusions on this aspect would require simultaneous (virtual and physical) data collection of a kind that is difficult to obtain by using purely passive means. Nonetheless, the insight obtained through analysis of the session logs about extent of frequency engagement, the timing and duration of such engagement, and its connection to characteristics of the player sheds light on this phenomenon of growing significance to society.

REFERENCES

1. Giuliano, G. Information Technology, Work Pattern and Intra-metropolitan Location: A Case Study. *Urban Studies* Vol. 35, No. 7, 1998, pp. 1077–1095.
2. Brewer, A. M., and D. Hensher. Distributed Work and Travel Behaviour: The Dynamics of Interactive Agency Choices Between Employers and Employees. *Transportation,* Vol. 27, No. 1, 2004, pp. 117–148.
3. Golob, T. F., and A. C. Regan. Impact of Information Technology on Personal Travel and Commercial Vehicle Operations: Research Challenges and Opportunities. *Transportation Research C,* Vol. 9, No. 2, 2001, pp. 87–121.
4. Mahmassani, H. S., J.-R. Yen, R. Herman, and M. Sullivan. Employee Attitudes and Stated Preferences Towards Telecommuting: An Exploratory Analysis. In *Transportation Research Record 1413,* TRB, National Research Council, Washington, D.C., 1993, pp. 31–41.
5. Mokhtarian, P. L., G. O. Collantes, and C. Gertz. Telecommuting, Residential Location, and Commute Distance Traveled: Evidence from State of California Employees. *Environment and Planning A,* Vol. 36, No. 10, 2004, pp. 1877–1897.
6. Mokhtarian, P. L. Telecommunications and Travel: The Case for Complementarity. *Journal of Industrial Ecology,* Vol. 6, No. 2, 2003, pp. 43–57.
7. Gould, J., and T. F. Golob. Consumer E-Commerce, Virtual Accessibility and Sustainable Transport. In *Social Change and Sustainable Transport* (W. R. Black and P. Nijkamp eds.), Indiana University Press, Bloomington, 2002.
8. Huang, Y., M. Zhu, J. Wang, N. Pathak, C. Shen, B. Keegan, D. Williams, and N. Contractor. The Formation of Task-Oriented Groups: Exploring Combat Activities in Online Games. *Proc., IEEE SocialCom-09,* Vancouver, British Columbia, Canada, 2009.
9. Williams, D., N. Martins, M. Consalvo, and J. Ivory. The Virtual Census: Representations of Gender, Race and Age in Video Games. *New Media and Society,* Vol. 11, No. 5, 2009, pp. 815–834.
10. Tychsen, A., M. Hitchens, and A. Drachen. Modeling and Analyzing Time in Multiplayer and Massively Multiplayer Games. *Games and Culture,* Vol. 4, No. 2, 2009, pp. 170–201.
11. Williams, D., N. Yee, and S. Caplan. Who Plays, How Much, and Why? Debunking the Stereotypical Gamer Profile. *Journal of Computer Mediated Communication* Vol. 13, No. 4, 2008, pp. 993–1018.
12. Chambers, C., W.-C. Feng, S. Sahu, and D. Saha. Measurement-Based Characterization of a Collection of On-Line Games. *Proc., Internet Measurement Conference,* 2005, pp. 1–14.
13. Huang, Y., C. Shen, D. Williams, and N. Contractor. Virtually There: The Role of Proximity and Homophily in Virtual World Networks. *Proc., IEEE SocialCom-09,* Vancouver, British Columbia, Canada, 2009.
14. Greene, W. H. *Econometric Analysis,* 5th ed., Prentice-Hall, Englewood Cliffs, N.J., 2003.

The Telecommunications and Travel Behavior Committee peer-reviewed this paper.

Cross-Clustered Model of Frequency of Home-Based Work Participation in Traditional Off-Work Hours

Erika Spissu, Naveen Eluru, Ipek N. Sener, Chandra R. Bhat, and Italo Meloni

A study was done to shed light on the determinants of working from home beyond the traditional office-based work hours. The frequency of work participation from home was examined for individuals who also have a traditional work pattern of traveling to an out-of-home workplace and a fixed number of work hours at the out-of-home workplace. The sample for the empirical analysis was drawn from the 2002 to 2003 Turin, Italy, survey of time use, which was designed and administered by the Italian National Institute of Statistics. The methodology recognizes both spatial and social clustering effects by using a cross-clustered ordered response structure to analyze the frequency of work participation from home during off-work periods. The model is estimated through the use of the inference technique of composite marginal likelihood, which represents a conceptually, pedagogically, and implementationally simpler procedure relative to traditional frequentist and Bayesian simulation techniques.

Rapid advances in information and communication technologies (ICTs) have substantially altered work patterns across the globe. Several studies have indicated that a consequence of the pervasiveness of the Internet is a blurring of the traditional separation between work and nonwork locations for conducting work (*1, 2*). A 2008 survey of 2,252 adult Americans reported that 19% of respondents had increased the amount of time spent working from home because of the availability of the Internet (*3*). As further evidence of the growing trend of teleworking from home, about 15% of U.S. workers worked remotely at home at least once a week in 2006 (*4*), and about 20% of European workers reported working at least a quarter of their working hours from home in 2005 (*5*).

Advances in ICTs are changing work not only in the context of space (where work is pursued) but in the context of time (when work is pursued). Studies in the social sciences and work habits literature (*6*) suggest that although ICTs provide a convenient means of obtaining and absorbing information almost instantaneously, they have also fed a workaholic culture, because people can work virtually any time but with consequent societal issues such as reduction of and interruptions to family time. There has long been recognition in the time use and activity literature of the important potential effects of ICT-related work patterns on individual time use and activity-travel patterns (*7–11*). These studies emphasized the importance of understanding work patterns as a precursor to generating and scheduling overall individual and household work and nonwork patterns.

Work patterns clearly have a role in shaping the way people conduct their day-to-day life, in general, and pursue their activity-travel patterns, in particular. However, the focus of earlier studies was on work location rather than the temporal dimension of work. This latter dimension typically is considered for the traditional arrangement of individuals who travel out of the home to work but is not considered for work patterns that entail working partly from home and partly from work. The emphasis of this paper is on the latter type of work pattern. Specifically, the frequency of work participation from home is examined for individuals who have the traditional work pattern of traveling to an out-of-home workplace and a fixed number of work hours at the out-of-home workplace. The data for the analysis are drawn from a time use survey conducted in Italy, where it is still rare that employees may telecommute or work away from their workplace during regular working hours. However, an increasing number of Italians are working from home outside traditional work hours, according to a research conducted in Turin, Italy (*12*).

The methodology was an ordered response system used to model frequency of work participation from home during traditionally off-work hours, explicitly recognizing both spatial and social clustering effects by using a multilevel structure. This is important since there may be unobserved effects (that is, those effects that cannot be directly captured through explanatory variables) based on spatial grouping effects (e.g., individuals residing in a certain neighborhood may be uniformly more likely to work off-hours due to spatial proximity effects) or on social grouping effects (e.g., individuals who interact closely with one another in social circles may be observed to cluster on the propensity to work off-hours from home; note that social grouping does not require any kind of spatial proximity). In such a multilevel clustering context, it is important to recognize and preserve between-cluster heterogeneity [i.e., intrinsic differences across clusters (*13, 14*)], because ignoring such heterogeneity, when present, would result in misestimated standard errors in linear models and inconsistent parameter estimation in nonlinear models. One also must consider local cluster-based variations in the relationship between the dependent and independent variables to avoid structural instability, especially in nonlinear models. Finally, heterogeneity among aggregate cluster units (neighborhoods or social groups) and heterogeneity among elementary units (individuals) must be differentiated. Ignoring

E. Spissu and Italo Meloni, CRiMM–Dipartimento di Ingegneria del Territorio, University of Cagliari, Via San Giorgio 12, 09124 Cagliari, Italy. N. Eluru, I. N. Sener, and C. R. Bhat, Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, 1 University Station, C1761, Austin, TX 78712-0278. Corresponding author: C. R. Bhat, bhat@mail.utexas.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 138–146.
DOI: 10.3141/2157-17

this differentiation and modeling the behavior of interest at a single level invites the pitfall of either the ecological fallacy, by which the level of analysis is solely at the aggregate level (i.e., failing to recognize that the elementary units act, not the aggregate units), or the atomistic fallacy, by which the analysis is pursued entirely at the elementary unit level (i.e., missing the context in which elementary units behave).

There has been substantial interest in multilevel analysis in several fields, including education, sociology, medicine, and geography. [Reise and Duan provide a review of multilevel models and their applications (*15*).] However, application of the method has been almost exclusively confined to the case of a strictly hierarchical clustering structure. This can be easily handled with a multilevel structure by including a random-effects term specific to each cluster and estimating the parameters of the resulting model by using the familiar maximum likelihood estimation (*16*). The situation changes when elementary units can be classified into more than one higher-level unit (more on this in the next section). The net result of such cross clustering is that the dimensionality of integration in the cross random-effects case explodes rapidly, making the likelihood maximization approach ineffective (*14*).

This paper adopts the technique of composite marginal likelihood (CML) estimation, an emerging inference approach in the statistics field, although there has been little to no coverage of this method in econometrics and other fields (*17, 18*). CML estimation is a simple approach that can be used when the full likelihood function is nearly impossible or is infeasible to evaluate because of underlying complex dependencies, as with econometric models with general cross-random effects. The CML approach also represents a conceptually, pedagogically, and implementationally simpler procedure relative to simulation techniques, and it has the advantage of reproducibility of results.

CROSS-LEVEL CLUSTERING APPROACHES

Strictly hierarchical multilevel analysis has seen substantial application, especially in the context of linear models. However, in the past decade multilevel analysis also has been applied to nonlinear models in the activity analysis field (*16, 19*). In the work of Bhat and Zhao (*16*), the clustering is purely spatial and is based on residential zone. These authors examined the number of daily shopping stops made by households while considering spatial clustering effects. In the work of Dugundji and Walker (*19*), the clustering is based on a combination of residential district or post code (to represent spatial clustering effects) and socioeconomic grouping (to proxy social interaction effects). These authors examined mode choice to work while accommodating spatial and social clustering effects. However, both these studies adopted strictly hierarchical clustering structures, wherein each individual is assigned to one and only one cluster, and the clusters are mutually exclusive and collectively exhaustive. Such structures lend themselves to maximum simulated likelihood estimation, since the strictly hierarchical clustering is accommodated through cluster-specific mixing random effects and individuals can be grouped into one of several clusters. The important point is that the dimensionality of integration of the probability expressions appearing in the Bhat–Zhao and Dugundji–Walker studies is independent of the number of clusters.

The only earlier study in the travel demand literature that captures cross-cluster effects is that by Bhat (*14*). Bhat, like Dugundji and Walker, models work mode choice, but he allows cross clustering based on residential location and work location. To allow maximum likelihood estimation, Bhat uses very aggregate spatial definitions of the work location, which reduces the dimensionality of the integration in the likelihood function and allows the use of simulation techniques. However, Bhat's simulation approach is infeasible in the more-general case of cross-cluster effects with several clusters in both dimensions, or when the cross-cluster effects are based on clustering in more than two dimensions. The main problem in these more general cases is that the dimensionality of integration is no more independent of the number of clusters in each dimension. To give a sense of the dimensionality, if Bhat had used the same spatial resolution of traffic analysis zones in defining work locations as in defining residential locations (193 traffic analysis zones), the number of dimensions of integration would have been of the order of 193 $*$ number of travel modes or 600 dimensions. This integration would have to be undertaken in Bhat's study over a conditional likelihood function integrand involving the product of the probabilities of each individual in the entire sample. Consequently, the likelihood maximization involves likelihood evaluations with numerically extremely small values, causing substantial instability problems. [Note that taking the logarithm of the likelihood function of the entire sample, as is the norm in the maximum likelihood method, offers no benefit because the log likelihood function does not simplify to the sum of the logarithm of the likelihood function of clusters involving fewer individuals than the entire sample. Even Bayesian techniques are impractical for the case of cross-random effects because they require extensive simulation and are time-consuming (*20*). In this regard, both the maximum likelihood and the Bayesian approach may be viewed as "brute force" simulation techniques that are not straightforward to implement and can create convergence assessment problems.]

The present paper applies a CML approach for cross clustering in the context of an ordered response structure. The CML approach, originally proposed by Lindsay (*21*), develops a surrogate likelihood function that involves easy-to-compute, low-dimensional, marginal likelihoods (*17, 18*). The CML approach is implemented here on the basis of the marginal likelihood of pairs of individuals. The approach is ideally suited for cross-random effects since it entails only bivariate distribution function evaluations, independent of the number of dimensions of clustering or the number of clusters within each dimension. [Bellio and Varin considered the CML approach for cross-random effects in generalized linear mixed models (*20*).] The CML approach can be applied by using simple optimization software for likelihood estimation and is based on a classical frequentist approach. Its basis in the theory of estimating equations (*21, 22*) ensures that the CML estimator is consistent, unbiased, and asymptotically normally distributed. The CML estimator (theoretically speaking) loses some efficiency relative to traditional maximum likelihood estimation, although this efficiency loss has been shown to be negligible in practice (*23*). In any case, the CML estimator is perhaps the only current practical approach for estimating parameters in general cross-random effects contexts.

METHODOLOGY

Model Structure

The model structure and estimation methodology are described in the general context of an ordered response model with two-dimensional cross-random clustering. In the substantive context of the current

paper, the dependent variable in the ordered response model corresponds to the frequency of work participation from home for individuals who have the traditional work pattern of traveling to an out-of-home workplace and have a fixed number of work hours at the out-of-home workplace. The two-dimensional clustering corresponds to spatial clustering based on the residential location of the individual and social clustering based on the social grouping to which the individual belongs. The specific manner in which the spatial and social clusters are defined and implemented in the empirical analysis is discussed subsequently.

In the usual framework of an ordered response model, let the underlying latent continuous random propensity z^*_{qij} of individual q in spatial cluster i and social cluster j be related to a vector x_{qij} of relevant explanatory variables as follows:

$$z^*_{qij} = \alpha_{ij} + \beta' x_{qij} + \epsilon_{qij}$$
$$z_{qij} = k \qquad \text{if } \psi_k < z^*_{qij} \le \psi_{k+1} \tag{1}$$

where

α_{ij} = scalar term associated with spatial group i and social group j,
β = vector of coefficients to be estimated,
ϵ_{qij} = standard normally distributed random term,
z_{qij} = observed ordinal frequency of working from home during off-work times,
k = index for the ordinal frequency category ($k = 1, 2, \ldots, K$), and
ψ_k = lower bound threshold for ordinal level k ($\psi_0 < \psi_1 < \psi_2 < \cdots < \psi_K < \psi_{K+1}$; $\psi_0 = -\infty$, $\psi_{K+1} = +\infty$).

ϵ_{qij} is assumed to be independent of the elements in α_{ij}, β, and x_{qij}. Also, as formulated here, x_{qij} does not include a constant term. The variance in the scalar term α_{ij} represents intrinsic unobserved heterogeneity across individuals in their propensity to work off-hours from home based on their residential location and social grouping.

Equation 1 represents the microlevel model for individuals. The scalar term α_{ij} is now allowed to vary across spatial clusters and social groups in a higher-level macro model:

$$\alpha_{ij} = \lambda' w_{ij} + u_i + v_j \tag{2}$$

where

w_{ij} = vector of observed variables specific to spatial cluster i or social group j or to the combination of spatial cluster i and social group j,
λ = parameter vector to be estimated, and
u_i, v_j = random terms that capture unobserved variations across spatial groups and social groups, respectively, in the propensity of working from home during off-work hours.

The latter two error terms are assumed to be realizations from independently and identically normal distributed terms across spatial and social clusters, respectively: $u_i \sim N(0, \sigma^2)$ and $v_j \sim N(0, \eta^2)$. Next, define $\gamma = (\beta', \lambda')'$ and $s_{qij} = (x'_{qij}, w'_{ij})'$. Then, the micro- and macromodels of Equations 1 and 2 can be combined to form

$$z^*_{qij} = \gamma' s_{qij} + u_i + v_j + \epsilon_{qij}$$
$$z_{qij} = k \qquad \text{if } \psi_k < z^*_{qij} \le \psi_{k+1} \tag{3}$$

The usual independence assumption of all error terms is invoked. If σ^2 (variance of u_i) and η^2 (variance of v_j) are equal to zero, then it is implied that there are no variations (due to unobserved factors) in the propensity to work off-hours from home across spatial and social clusters, respectively. In this case, the cross-random ordered response (CROR) model of Equation 3 collapses to the standard ordered response (SOR) model. The implication is that all unobserved heterogeneity is due to interindividual differences, and there is no unobserved heterogeneity based on spatial and social clustering. The propensity specification of Equation 3 generates a covariance pattern among individuals as follows: for two individuals in the same spatial cluster, but not in the same social cluster, the covariance in their propensities to work off-hours is σ^2. For two individuals in the same social cluster, but not in the same spatial cluster, the covariance is η^2. For two individuals in the same spatial and social cluster, the covariance is $\sigma^2 + \eta^2$. Finally, for two individuals not in the same spatial cluster or in the same social cluster, the covariance is zero.

Estimation Approach

A pairwise marginal likelihood estimation approach is used here, which corresponds to a composite marginal approach based on bivariate margins (*18, 24–26*). Each bivariate margin represents the joint probability of the observed frequency of working off-hours from home for a pair of individuals q and h in the sample. The presence of spatial and social clustering effects leads to covariance effects between the pair of individuals q and h based on their spatial and social groupings.

In this section, since each individual q is uniquely identified with a particular spatial cluster i and a particular social cluster j, it is convenient for presentation to suppress the indices i and j. Thus, let z_q stand for z_{qij} and s_q for s_{qij}. Also, let d_q be the actual observed ordinal frequency of working from home during off-work hours for individual q. The pairwise marginal likelihood function may then be written, after defining $\delta = (\gamma', \sigma, \eta)'$, as

$$L_{\text{CML}}(\delta) = \prod_{q=1}^{Q-1} \prod_{h=q+1}^{Q} \left[P(z_q = d_q, z_h = d_h) \right]$$

$$P(z_q = d_q, z_h = d_h) = \Phi_2\left(u_{(d_q)}, u_{(d_h)}, \theta_{qh}\right) - \Phi_2\left(u_{(d_q-1)}, u_{(d_h)}, \theta_{qh}\right) - \Phi_2\left(u_{(d_q)}, u_{(d_h-1)}, \theta_{qh}\right) + \Phi_2\left(u_{(d_q-1)}, u_{(d_h-1)}, \theta_{qh}\right)$$

$$u_{(d_q)} = \left(\frac{\psi_{(d_q)} - \gamma' s_q}{\sqrt{1 + \sigma^2 + \eta^2}} \right)$$

and

$$\theta_{qh} = \frac{G_{qh}\sigma^2 + R_{qh}\eta^2}{1 + \sigma^2 + \eta^2} \tag{4}$$

$G_{qh} = 1$ if q and h are in same spatial cluster and $G_{qh} = 0$ otherwise. Similarly, $R_{qh} = 1$ if q and h are in same social cluster and $R_{qh} = 0$ otherwise. The bivariate probability expressions in the pairwise marginal likelihood function of Equation 4 are straightforward to compute, since they entail only four bivariate standard normal expressions. The pairwise marginal likelihood function comprises $Q(Q - 1)/2$ pairs of bivariate probability computations, which can itself become quite time-consuming. Fortunately, the individuals

that have no spatial and no social interdependencies can be preidentified. The coding exploits this situation to enable relatively fast maximization of the logarithm of the pairwise marginal likelihood function.

The CML estimator obtained by maximizing the logarithm of the function in Equation 4 with respect to the γ, σ, and η parameters is consistent and asymptotically normal distributed with the asymptotic variance matrix given by Godambe's sandwich information matrix (*27*):

$$F(\delta) = \frac{1}{Q}[H(\delta)]^{-1} J(\delta)[H(\delta)]^{-1} \tag{5}$$

where

$$H(\delta) = E\left[-\frac{\partial^2 \log L_{\text{CML}}(\delta)}{\partial\delta\partial\delta'}\right]$$

and

$$J(\delta) = E\left[\left(\frac{\partial \log L_{\text{CML}}(\delta)}{\partial\delta}\right)\left(\frac{\partial \log L_{\text{CML}}(\delta)}{\partial\delta'}\right)\right]$$

The "bread" matrix $H(\delta)$ of Equation 5 can be estimated straightforwardly by using the Hessian of the negative of log $L_{\text{CML}}(\delta)$, evaluated at the CML estimate $\hat{\delta}$. This is because the information identity remains valid for each pairwise term forming the composite marginal likelihood. Thus, $H(\delta)$ can be estimated as

$$\hat{H}(\hat{\delta}) = \left(-\left[\sum_{q=1}^{Q-1}\sum_{h=q+1}^{Q} \frac{\partial^2 \log L_{\text{CML},qh}(\hat{\delta})}{\partial\hat{\delta}\partial\hat{\delta}'}\right]\right) \tag{6}$$

where

$$L_{\text{CML},qh}(\hat{\delta}) = [P(z_q = d_q, z_h = d_h)] \mid \hat{\delta}$$

However, the estimation of the "vegetable" matrix $J(\delta)$ is more difficult, since $\partial\log L_{\text{CML}}(\delta)/\partial\delta$ vanishes when evaluated at the CML estimate $\hat{\delta}$. Further, one cannot estimate $J(\delta)$ as the sampling variance of individual contributions to the composite score function because of the underlying spatial and social dependence in observations. In addition, the nondecaying correlation pattern of the current framework does not permit the use of the windows resampling procedure of Heagerty and Lumley (*28*) to estimate $J(\delta)$ as in Bhat et al. (*18*). Hence pure Monte Carlo computation is used to estimate $J(\delta)$(*18*). In this approach, B data sets $Z^1, Z^2, \ldots, Z^B$ are generated, where Z^b ($b = 1, 2, \ldots, B$) is a vector of one possible realization for $(z_1, z_2, \ldots, z_Q)$ for the exogenous variable vector $S = (s_1, s_2, \ldots, s_Q)$ (under the assumed model with $\delta = \hat{\delta}(S)$). Once these data sets are generated, the estimate of $J(\delta)$ is given by

$$J(\delta) = \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\partial \log L_{\text{CML}}(\delta)}{\partial\delta}\right)\bigg|_{Z^b}\left(\frac{\partial \log L_{\text{CML}}(\delta)}{\partial\delta'}\right)\bigg|_{Z^b}\right]$$

This computation is not very demanding because the model in Equation 1 can be generated in a straightforward manner. Various values of B were tested for stability in the estimate of $J(\delta)$, and it was found that a value of $B = 1000$ was more than adequate for reasonable accuracy.

DATA AND SPATIAL–SOCIAL CLUSTER DEFINITIONS

Data Sources and Sample Used

The primary source of data used in this paper is the 2002 to 2003 Turin time use survey, which was designed and administered by the Italian National Institute of Statistics (ISTAT) and sponsored by the Turin town council and 14 neighboring town councils (Baldissero Torinese, Beinasco, Borgaro Torinese, Collegno, Grugliasco, Moncalieri, Nichelino, Orbassano, Pecetto Torinese, Pino Torinese, Rivoli, San Mauro Torinese, Settimo Torinese, and Venaria Reale). The survey collected a daily-activity time use diary from each of 4,537 household members age 3 years and older from 1,830 households (*29*). Detailed work characteristics and demographic information were obtained from all surveyed households and individuals, including residence location in one of the 24 municipalities in the Turin area. [Information regarding household residence within the 10 districts of Turin city, although collected in the time use survey, was missing in the data. However, the authors had access to the set of sampled households, with household and individual characteristics, in each municipality (the location data set). As a means to identify each sampled household in the time use survey with a specific municipality in the location data set, a probabilistic linking procedure based on individual and household demographics was sued. A household in the time use survey was linked to a household in the location data set only if the probability of a true match was 0.95 or higher. The details of this linking procedure are available on request from the authors.]

In addition to the main time use survey data, the data sets of the 2001 census of population (*30*) and the 2001 census of industry and services (*31*) were used to obtain built-environment variables, including housing type measures; number of commercial, industrial, and service units; population density; household density; and employment density.

The final sample used for the estimation includes 2,042 individuals age 14 and older. Since the focus of the study is work participation during traditionally off-work hours, only workers (employed individuals as identified in the survey) were considered in the analysis. The survey specifically asked workers to provide the frequency of off-hours work participation from home in ordered response categories, which serves as the dependent variable for the model. These categories and the corresponding distribution among workers are as follows: never, 1,597 respondents (78.2%); few times a month, 173 respondents (8.5%); few times a week, 182 respondents (8.9%); and everyday, 90 respondents (4.4%). These sample shares of individuals working during off-hours at home, as obtained from the Turin area, are similar to corresponding figures for all of Italy (*29*).

Spatial–Social Cluster Definitions

As discussed earlier, this study incorporates both spatial and social clustering effects. An issue in introducing such effects is that the analyst has to define the spatial unit and the social unit for accommodating the clustering effects. The present study used the finest spatial resolution to which individual residential locations could be identified with, which corresponded with locations in 24 municipalities (14 municipalities outside Turin and 10 municipalities within Turin). Two different spatial clustering schemes were considered based on whether workers were resident in the same zone and whether workers were resident in the same or immediately

adjoining zone. The second spatial clustering scheme is more expansive than the first.

A number of social clustering schemes may be defined based on the social circle in which each worker moves as part of her or his daily life. In the social science literature, various characterizations of social interaction and social proximity have been proposed (*32*) based on direct relationships (such as with family members, work colleagues, and friends), indirect relationships (with individuals living in the same environment), and cultural relationships (with individuals of the same social status, race, and ethnicity). In the transportation literature, social interactions and proximity effects have been examined in one of the following two broad ways: by using social network geographies (*33*) and by using the egocentric approach (*34*). The former approach focuses on identifying all the possible social connections of each individual through a survey or other data-collection method. The latter approach clusters individuals in social groups based on their demographics or attitudes toward joint activities. In the present analysis, there is no explicit information on the social circle of each individual in the sample, but there is rich demographic information on each individual. Thus, the definition of social interactions is based on an egocentric approach as in Dugundji and Walker (*19*), where social interactions are considered to be particularly strong among individuals who share certain common demographic characteristics. (Note, however, that the methodology proposed in this paper is generic and can be applied with any approach that identifies the social network of an individual.) The demographic groupings considered in the empirical analysis included the following key variables (and also groupings based on various possible combinations of these key variables): whether workers share the same marital status (never married, currently married, or separated, divorced, or widowed) and whether workers have children (including examination of children in different age groups).

Regardless of the definitions considered for the social clustering effect, there is always a diversity of individuals in each municipality based on social grouping. Thus, it is not possible to have independent and mutually exclusive clusterings of individuals in the sample based on both spatial and social clustering. That is, the cross-random effects lead to a global interaction network over the entire sample, leading to a kind of a giant cluster with dependencies between each pair of workers.

MODEL ESTIMATION RESULTS

Variable Specification and Cluster Definitions

Several different exogenous variables, functional forms of variables, and variable interactions were considered in the model specifications. The exogenous variables included (*a*) individual demographics (age, gender, etc.); (*b*) individual work-related characteristics (work schedule, number of jobs, etc.); (*c*) household characteristics (household composition, household income, etc.); and (*d*) built-environment measures of the individual's residence location. (Occupation status was collected in an ambiguous way that did not allow its use as an exogenous variable; inclusion of occupation status in the analysis is important in further research.) The final specification was based on a systematic process of removing statistically insignificant variables and combining variables when their effects were not significantly different. The specification process was also guided by previous research and intuitiveness and parsimony considerations. For the continuous variables in the data (such as age, work hours at the out-of-home work location, and income), alternative functional forms were tested that included a linear form, a spline (or piecewise linear) form, and dummy variables for different ranges. For all the continuous variables, dummy variables provided the best results and are used in the final specification (income, however, did not turn out to be statistically significant). Various threshold values for defining the dummy variables were tested, and those that provided the best fit were used.

In addition to several variable specifications, various spatial and social clustering schemes were considered through the specification of the G_{qh} and R_{qh} dummy variables. The best specification was obtained with a spatial clustering scheme based on whether two workers reside in the same municipality and the social clustering scheme based on whether workers share marital status (never married, married, or separated, divorced, or widowed). The following discussions present only the results of the final variable specification with the best spatial or social clustering scheme.

Estimation Results

Table 1 presents the model estimation results. The base category is listed in parenthesis for each discrete exogenous variable. The

TABLE 1 CROR Estimation Results

Characteristic	CROR Estimates	CROR Est./SE
Individual Characteristics		
Age (>45 is the base)		
14–30	−0.410	−3.87
31–45	−0.202	−2.52
Education (<2–3 years of high school is the base)		
High-level education (bachelor's degree or higher)	1.146	11.52
Medium-level education (4–5 years of high school and a bachelor's degree)	0.673	8.46
Work Characteristics		
Work hours per week (40 h and less is the base)		
41 h and more	0.314	4.64
Work schedule involves working in the		
Evening	0.207	2.38
Night	−0.340	−2.60
Saturday	0.383	4.92
Having a second job or occupation (not self-employed is the base)	0.555	3.66
Self-employed	0.331	3.91
Household Characteristics		
Presence of young adults 18–24 years (presence of young children or no children is the base)	−0.230	−2.02
Availability of Internet access at home	0.147	2.21
Presence of a housekeeper	0.392	3.90
Thresh01	1.290	9.02
Thresh02	1.712	11.74
Thresh03	2.454	16.31
Correlation Terms		
Spatial correlation (σ)	0.114	2.71
Social correlation (η)	0.061	1.75
Mean log likelihood	−1,321.650	

coefficients in the table indicate the effects of variables on the latent propensity of work participation from home during off-work hours (that is, they represent elements of the γ vector as defined in the section on model structure). Although many functional forms were tested for several variables, the final specification includes only dummy variables. Thus, the magnitudes of the coefficients provide an estimate of the importance of the variables in influencing off-work participation propensities and participation probabilities.

Individual Demographics

The effects of individual demographics indicate that young individuals, <45 years, are less likely than individuals older than 45 to work from home during off-work hours. This is particularly the case for the very young individuals (≤30). This may reflect the generally wider social networks of young individuals (*35*) and the generally higher participation rates of younger individuals in out-of-home discretionary and maintenance activities (*36–39*). Such tendencies would reduce the time available for, or the inclination to, work at home outside standard work hours. Further, the results point to the increased home-based work during off-work times among more educated individuals relative to less-educated individuals, perhaps because of demanding and high-status work positions.

Work Characteristics

Among the work-related characteristics, the results suggest that individuals who work more than 40 h a week at their out-of-home workplace are also more likely to work during off-work hours, presumably because of overall "workaholic" tendencies. The work schedule variables show that individuals who work in the evening and on Saturdays are more likely to bring work home compared to other workers, while those who work at night are less likely to bring work home relative to those who work at any other time. In addition, individuals with a second job are more likely to bring work home, probably because of time pressures and the need to juggle multiple things at multiple jobs. Self-employed individuals are more likely to work during off-work times relative to those who are not self-employed. (In Italy, it is the norm for self-employed individuals to have an out-of-home office location, with set times of work at the out-of-home office.) This is not surprising, since self-employed individuals have an incentive to work harder and longer to ensure that their venture is successful.

Household Characteristics

Among household characteristics, the results show a lower tendency to bring work home among individuals who live in households with young adults ages 18 to 24 years (i.e., driving license age) compared to individuals living in households with young or no children. This result needs further exploration, but a possible reason for it is more social interaction within the family when there are multiple adults in the household (relative to when there are young children or no children). As expected, Internet access at home increases the tendency to work from home during off-work hours (*6*). This raises the issue of whether the presence of Internet access makes individuals work more from home during off-work times or whether the need to work during off-work time motivates individuals to set an up Internet connection at home. Although both effects are possible (and may apply to different sets of individuals), Internet access is becoming a way of life for maintaining social connections and family connectedness (*3*). Thus, it is likely that having an Internet connection at home is not solely influenced by a desire to work from home during off-work times, but, once available, Internet access draws people into work-related activities. Finally, the presence of a paid housekeeper has a positive effect on the propensity to work from home during off-work times. A household contracting out household tasks is a possible sign of work spillover effects at home.

Built-Environment Attributes of Residence Location

Several built-environment measures associated with the municipality of each individual's residence were considered in the model specifications. However, none of these were statistically significant, possibly because rather aggregate spatial units were used to compute built-environment measures. Future research efforts should explore finer resolutions of space for computing built-environment measures.

Threshold Parameters

The thresholds do not have substantial interpretations. They simply serve to translate the underlying latent propensity to work off-hours to the ordinal categories of working off-hours.

Effects of Spatial and Social Dependence

The CROR model accommodates spatial dependence among workers residing in the same municipality and social dependence among workers based on marital status. The variance terms σ^2 (variance of u_i in Equation 3) and η^2 (variance of v_j in Equation 3) provide the magnitude of this dependence, as indicated in the section on model structure. Table 1 indicates the estimated standard deviations σ and η, both of which are statistically significant at the 0.05 level of significance. (Note that the standard deviation terms should be positive, so a one-tailed *t*-test is warranted.) The spatial clustering effect is particularly strong and highly statistically significant. For the CML log likelihood value at convergence, the CROR value is −2698809, compared to the SOR value of −2700118. The composite marginal likelihood ratio test (CLRT) statistic, computed as twice the difference in the composite marginal log likelihood values, yields a value of 2618. However, this CLRT statistics does not have the standard chi-squared asymptotic distribution under the null hypothesis as in the case of the maximum likelihood inference procedure. Although one can use a bootstrapping approach to obtain the precise distribution of the CLRT statistic, several bootstrapping runs are needed, which becomes cumbersome. In any case, the *t*-statistics on the σ and η parameter estimates are statistically significant, which indicates the data fit superiority of the CROR model over the SOR model. Further, as in the usual likelihood approach, one may compute an adjusted rho-bar-squared value $\bar{\rho}_c^2$ in the composite marginal likelihood approach for the CROR model and the SOR models as $\bar{\rho}_c^2 = 1 - [(\log L_{\text{CML}}(\hat{\delta}) - H)/\log L_{\text{CML}}(C)]$, where $\log L_{\text{CML}}(\hat{\delta})$ is the composite marginal log likelihood at convergence, H is the number of model parameters excluding the thresholds, and $\log L_{\text{CML}}(C)$ is the log likelihood with only thresholds in the model. The value of $\bar{\rho}_c^2$ for the CROR model is 0.142, and that for the SOR model is 0.141.

Aggregate Elasticity Effects

The parameters on the exogenous variables in Table 1 provide a sense of the relative magnitudes of effects of variables on the propensity to work from home during off-work hours. This is because all variables in the model are dummy variables. The results indicate that the education variables and having a second job have the highest impact, followed by presence of a housekeeper, young age (between 14 and 30 years), characteristics of work schedule, self-employment status, and hours of work at the out-of-home work location. Internet access does not have as substantial an effect as the other variables.

The coefficients in Table 1, although providing a sense of relative magnitudes of the effect of variables on the propensity to work from home during off-work hours, do not provide the absolute magnitude of effects of variables on the probability of each frequency level of working from home during off-work periods. To obtain such absolute magnitude effects, the aggregate-level elasticity effects of variables are computed. In particular, the value of each dummy explanatory variable is changed to 1 for the subsample of observations for which the variable takes a value of 0 and changed to 0 for the subsample of observations for which the variable takes a value of 1. The shifts are then summed in expected aggregate shares of each frequency level of working from home in the two subsamples after the sign of the shifts in the second subsample is reversed and an effective percentage change in the expected aggregate share of teenagers participating in each number of activity episodes due to a change in the dummy variable from 0 to 1 is computed.

The elasticity effects are presented in Table 2 for each frequency level of working from home during off-work hours. The numbers in the table may be interpreted as the percentage change in the probability of working from home due to a change in the variable from 0 to 1. For instance, the first number in Table 2 indicates that the probability of a young individual (14 to 30 years old) never working from home is 11.55% higher than the probability of an individual older than 45 years never bringing work home, other characteristics being equal. The results in Table 2 confirm the relative magnitude effects of variables discussed earlier. Note that in Italy, having Internet access is trumped by other factors in the effect on working from home beyond the usual workday at the office. Recent studies have focused on how technology affects home-based work and work–life balance (*2, 40*), which is an important and valuable direction of research. However, it appears from the results of the present study that work characteristics and demographics of individuals and households play a much more important role in determining who works from home during off-work hours and who does not. In particular, those with high education levels, who work multiple jobs, and are self-employed are particularly predisposed to working from home.

CONCLUSIONS

Rapid advances in telecommunications and information technologies (such as computers and smart phones) have altered work patterns across the globe. These advances affect work not only in the context of space (i.e., where work is pursued) but in the context of time (i.e., when work is pursued). Emphasis in the work-based research literature has been on the spatial context—telecommuting from home, working at a satellite work location, working at the regular

TABLE 2 Aggregate Elasticity Effects

	Propensity to Work from Home During Off-Work Hours			
Characteristic	Never	Few Times a Month	Few Times a Week	Every Day
Individual Characteristics				
Age (>45 is the base)				
14–30	11.55	−32.41	−43.40	−58.73
31–45	6.10	−15.90	−22.63	−34.04
Education (<2–3 years of high school is the base)				
High-level education (bachelor's degree or higher)	−44.18	82.03	160.00	321.52
Medium-level education (4–5 years of high school and a bachelor's degree)	−20.83	52.09	74.57	126.38
Work Characteristics				
Work hours per week (40 h and less is the base)				
41 h and more	−10.16	26.19	38.23	56.22
Work schedule involves working in the				
Evening	−6.55	16.72	24.46	36.94
Night	9.38	−26.56	−34.96	−47.90
Saturday	−11.66	31.57	43.89	61.61
Having a second job or occupation (not self-employed is the base)	−19.68	42.11	71.46	131.58
Self-employed	−10.87	27.40	40.77	61.71
Household Characteristics				
Presence of young adults 18–24 years (presence of younger children or no children is the base)	6.61	−18.35	−24.68	−34.36
Availability of Internet access at home	−4.48	11.97	16.76	24.20
Presence of a housekeeper	−13.42	31.46	49.99	81.71

work location—rather than the temporal context of work. The time dimension is typically considered for the traditional arrangement of individuals who travel out of the home to work (work start time and work end time at the traditional workplace) but not for work patterns that include working partly from home and partly from work on a workday, or work patterns that involve working from home beyond the traditional office-based work hours. This study sought to understand the determinants of working from home beyond the traditional office-based work hours. The frequency of work participation from home for individuals who also have the traditional work pattern of traveling to an out-of-home workplace with a fixed number of work hours at the workplace was examined. The sample for the empirical analysis was drawn from the 2002 to 2003 Turin time use survey, designed and administered by ISTAT.

The methodology recognized both spatial and social clustering effects by using a cross-clustered ordered response structure. This is important since there may be unobserved effects (that is, those effects that cannot be directly captured through explanatory variables) based on spatial or social grouping effects. The net result of such cross clustering is that the dimensionality of integration in the case of cross-random effects explodes rapidly, making the likelihood maximization approach ineffective if not infeasible. This study adopted CML estimation, which represents a conceptually, pedagogically, and implementationally simpler procedure relative to simulation techniques and has the advantage of reproducibility of results.

The results of the analysis indicate the statistically significant presence of spatial and social clustering in the underlying propensity to work from home during off-work hours. Ignoring such clustering can lead to inconsistent model parameter results and poorer data fit. The results also demonstrate the practical flexibility of the CML approach in situations in which traditional maximum likelihood or Bayesian techniques are either practically infeasible or ineffective. Several variables were considered in the model specifications as determinants of the frequency of home-based work during traditionally off-work hours. The results show that age, education level, number of hours of work, work schedule, number of jobs, self-employment status, presence of young adults in the household, Internet access at home, and presence of a housekeeper all have statistically significant effects. Among these, education level and number of jobs have the most dominant effects, whereas home Internet access has a relatively minor effect. None of the built-environment variables turned out to affect working from home during off-work periods, although this result needs further exploration in future studies with use of a fine resolution of space to compute built-environment measures.

The empirical results have implications for individual time use and activity-travel patterns, especially in the context of how home-based work during the normal off-work periods may alter the generation and scheduling of daily household work and nonwork patterns. In this regard, the model system developed in this paper can constitute one element of a larger-scale, activity-based travel demand modeling system that uses demographic and work characteristics (at the out-of-home workplace) to predict complete daily individual activity-travel patterns. The results from the current analysis should be useful in family and social science research, since research in those areas suggests that working from home affects individuals' satisfaction levels with family life, social interactions, and leisure participation. For instance, spending more time working at home can translate to less time with family, less time eating meals with family, less time socializing with neighbors and friends, and less time pursuing hobbies and other recreational pursuits.

Examining work-related activities at home and identifying those who pursue such work patterns is important both for predicting activity-travel patterns and because of the far-reaching effects on how people live and spend time. This is a relatively underresearched area, and it is hoped that this research will encourage focus on not only the where but also the when of work patterns. Although the Italian data set used in this analysis is appropriate in many ways, it is limited in the exogenous variables collected and available to predict the frequency of working from home during traditionally off-work times. Future efforts should consider a more comprehensive set of variables.

ACKNOWLEDGMENTS

The authors acknowledge the helpful comments of three anonymous reviewers on an earlier version of the paper. The authors thank the Turin Town Council Office of Statistics for providing data from the time use survey conducted in Turin. The corresponding author acknowledges the support of an international visiting professorship from the University of Cagliari. Funding from the Southwest Region University Transportation Center is also acknowledged. The authors thank Lisa Macias for her help in formatting this document.

REFERENCES

1. Sall, E. A., and C. R. Bhat. An Analysis of Weekend Work Activity Patterns in the San Francisco Bay Area. *Transportation,* Vol. 34, No. 2, 2007, pp. 161–175.
2. Valcour, P., and L. W. Hunter. Technology, Organizations, and Work-Life Integration. In *Work and Life Integration: Organizational, Cultural, and Individual Perspectives* (E. E. Kossek and S. J. Lambert, eds.), Erlbaum Associates, Mahwah, N.J., 2005, pp. 61–84.
3. Kennedy, T. L. M., A. Smith, A. T. Wells, and B. Wellman. *Networked Families.* Pew Internet and American Life Project. 2008. http://www.pewinternet.org/.
4. *Telework Trendlines for 2006.* WorldatWork. Feb. 2007. http://www.worldatwork.org/.
5. Fourth European Working Conditions Survey. European Foundation for the Improvement of Living and Working Conditions. 2007. http://www.eurofound.europa.eu/.
6. Kaufman-Scarborough, C. Time Use and the Impact of Technology Examining Workspaces in the Home. *Time and Society,* Vol. 15, No. 1, 2006, pp. 57–80.
7. Golob, T. F., and A. C. Regan. Impacts of Information Technology on Personal Travel and Commercial Vehicle Operations: Research Challenges and Opportunities. *Transportation Research Part C,* Vol. 9, No. 2, 2001, pp. 87–121.
8. Bailey, D. E., and N. B. Kurland. A Review of Telework Research: Findings, New Directions, and Lessons for the Study of Modern Work. *Journal of Organizational Behavior,* Vol. 23, No. 4, 2002, pp. 383–400.
9. Srinivasan, S., and C. R. Bhat. An Exploratory Analysis of Joint-Activity Participation Characteristics Using the American Time Use Survey. *Transportation,* Vol. 35, No. 3, 2008, pp. 301–328.
10. Mokhtarian, P. L., I. Salomon, and S. L. Handy. The Impacts of ICT on Leisure Activities and Travel: A Conceptual Exploration. *Transportation,* Vol. 33, No. 3, 2006, pp. 263–289.
11. Pinjari, A. R., and C. R. Bhat. A Multiple Discrete-Continuous Nested Extreme Value (MDCNEV) Model: Formulation and Application to Non-Worker Activity Time-Use and Timing Behavior on Weekdays. *Transportation Research Part B,* Vol. 44, No. 4, 2008, pp. 562–583.
12. Belloni, A. M. *Andare a tempo, Il caso di Torino: una ricerca sui tempi della città,* Franco Angeli, Milan, Italy, 2007.
13. Jones, K., and C. Duncan. People and Places: The Multilevel Model as a General Framework for the Quantitative Analysis of Geographical Data. In *Spatial Analysis: Modelling in a GIS Environment* (P. Longley and M. Batty, eds.), GeoInformation International, Cambridge, United Kingdom, 1996, pp. 79–104.

14. Bhat, C. R. A Multi-Level Cross-Classified Model for Discrete Response Variables. *Transportation Research Part B,* Vol. 34, No. 7, 2000, pp. 567–582.
15. Reise, S. P., and N. Duan (eds.). *Multilevel Modeling: Methodological Advances, Issues, and Applications.* Lawrence Erlbaum, Mahwah, N.J., 2003.
16. Bhat, C. R., and H. Zhao. The Spatial Analysis of Activity Stop Generation, *Transportation Research Part B,* Vol. 36, No. 6, 2002, pp. 557–575.
17. Varin, C. On Composite Marginal Likelihoods. *Advances in Statistical Analysis,* Vol. 92, No. 1, 2008, pp. 1–18.
18. Bhat, C. R., I. N. Sener, and N. Eluru. A Flexible Spatially Dependent Discrete Choice Model: Formulation and Application to Teenagers' Weekday Recreational Activity Participation. *Transportation Research Part B.* 2010. http://www.sciencedirect.com/science?_ob=ArticleListURL&_method=list&_ArticleListID=1342345305.
19. Dugundji, E. R., and J. L. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1921,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.
20. Bellio, R., and C. Varin. A Pairwise Likelihood Approach to Generalized Linear Models with Crossed Random Effects. *Statistical Modelling,* Vol. 5, No. 3, 2005, pp. 217–227.
21. Lindsay, B. G. Composite Likelihood Methods. *Contemporary Mathematics,* Vol. 80, 1988, pp. 221–239.
22. Cox, D., and N. Reid. A Note on Pseudolikelihood Constructed from Marginal Densities. *Biometrika,* Vol. 91, No. 3, 2004, pp. 729–737.
23. Varin, C., and P. Vidoni. Pairwise Likelihood Inference for General State Space Models. *Econometric Reviews,* Vol. 28, No. 1, 2009, pp. 170–185.
24. Varin, C., and C. Czado. *A Mixed Autoregressive Probit Model for Ordinal Longitudinal Data.* 2009. http://www-m4.ma.tum.de/Papers/Czado/varin_czado_revision7.pdf.
25. Apanasovich, T. V., D. Ruppert, J. R. Lupton, N. Popovic, N. D. Turner, R. S. Chapkin, and R. J. Carroll. Aberrant Crypt Foci and Semiparametric Modelling of Correlated Binary Data. *Biometrics,* Vol. 64, No. 2, 2008, pp. 490–500.
26. Ferdous, N., N. Eluru, C. R. Bhat, and I. Meloni. A Multivariate Ordered Response Model System for Adults' Weekday Activity Episode Generation by Activity Purpose and Social Context. *Transportation Research Part B.* 2010. http://www.sciencedirect.com/science?_ob=ArticleListURL&_method=list&_ArticleListID=1342412622.
27. Godambe, V. An Optimum Property of Regular Maximum Likelihood Estimation. *Annals of Mathematical Statistics,* Vol. 31, No. 4, 1960, pp. 1208–1211.
28. Heagerty, P. J., and T. Lumley. Window Subsampling of Estimating Functions with Application to Regression Models. *Journal of the American Statistical Association,* Vol. 95, 2000, pp. 197–211.
29. Istat. Multipurpose Time Use Survey. 2006. http://www.istat.it/.
30. Istat. Census of Population. 2001. http://www.istat.it/.
31. Istat. Census of Industry and Services. 2001. http://www.istat.it/.
32. Agrawal, A., D. Kapur, and J. McHale. How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data. *Journal of Urban Economics,* Vol. 64, No. 2, 2008, pp. 258–269.
33. Axhausen, K. W. Social Networks and Travel: Some Hypotheses. In *Social Dimension of Sustainable Transport: Transatlantic Perspectives* (K. P. Donaghy, S. Poppelreuter, and G. Rudinger, eds.), Ashgate, Aldershot, United Kingdom, 2005, pp. 90–108.
34. Carrasco, J. A., B. Hogan, B. Wellman, and E. J. Miller. Collecting Social Network Data to Study Social Activity-Travel Behavior: An Egocentric Approach. *Environment and Planning B: Planning and Design,* Vol. 35, No. 6, 2008, pp. 961–980.
35. Frei, A., and K. W. Axhausen. *Size and Structure of Social Network Geographies.* Arbeitsberichte Verkehrs und Raumplanung 439. IVT, ETH Zurich, Switzerland, 2007.
36. Chen, C., and P. Mokhtarian. Tradeoffs Between Time Allocations to Maintenance Activities/Travel and Discretionary Activities/Travel. *Transportation,* Vol. 33, No. 3, 2006, pp. 223–240.
37. Bhat, C. R. A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions. *Transportation Research Part B,* Vol. 39, No. 8, 2005, pp. 679–707.
38. Meloni, I., M. Bez, and E. Spissu. Random Utility-Based Microeconomic Model for Discretionary Activity Time Allocation. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2054,* Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 74–82.
39. Srinivasan, K. K., and S. R. Athuru. Analysis of Within-Household Effects and Between Household Differences in Maintenance Activity Allocation. *Transportation,* Vol. 32, No. 5, 2005, pp. 495–521.
40. Lewis, S., and C. L. Cooper. The Work-Family Research Agenda in Changing Contexts. *Journal of Occupational Health Psychology,* Vol. 4, No. 4, 1999, pp. 382–393.

The Telecommunications and Travel Behavior Committee peer-reviewed this paper.

Influence of E-Shopping on Shopping Travel

Evidence from Minnesota's Twin Cities

Xinyu Cao, Frank Douma, and Fay Cleaveland

Research was done to reveal the travel impact of e-shopping in the Minneapolis–St. Paul (Twin Cities) metropolitan area of Minnesota. A sample of Internet users drawn from urban, suburban, and exurban neighborhoods was used to identify the relationship between e-shopping and in-store shopping. An online survey composed of direct and attitudinal questions was used to obtain the data. Ordered probit models were developed to account for the influences of a variety of confounding factors, such as shopping attitudes, shopping accessibility, shopping responsibility, and sociodemographics. The preliminary results, controlled for the confounding factors, show that e-shopping behavior (for online searching and online buying) tends to have a complementarity effect on in-store shopping.

E-shopping has the potential to replace traditional in-store shopping. It is well perceived that information and communication technologies (ICTs) have pervasive effects on modern society—they are changing how and where people work, shop, and live life. Significant research has been conducted to understand the impact of ICTs on where work is done and how this affects travel. E-shopping has become a centerpiece of ICTs because of its proliferation. In the United States, online retail spending grew by 19% a year to \$136.4 billion in 2007, accounting for 4% of total retail sales (*1*). In the Netherlands, online sales have increased from less than €200 million in 1999 to more than €1.6 billion in 2004, as Farag reported (*2*). The growth of e-shopping has reshaped consumers' shopping behavior. Online buying could be a substitute for traditional shopping media and may well dominate the exchange of certain products (e.g., digital assets) in the future (*3*). E-shopping as used here refers to the business-to-consumer segment of e-commerce (*4*), that is, product information search (online searching) and product transactions (online buying) via Internet, unless otherwise indicated.

Transportation planners are interested in the changes that e-shopping and its long-term potential in the retail industry will bring to transportation systems, having a particular focus on the effects of e-shopping on individuals' activity-travel patterns and on freight transport. Changes in freight transport can result from the growth of delivery trips to consumers and the bypassing of wholesalers and retailers in the network from manufacturer to consumer (*5*). The potential of online buying to substitute for traditional in-store shopping and reduce personal shopping travel has important implications for travel demand management and congestion mitigation. According to the 2001 National Household Transportation Survey, on average, shopping travel accounted for 14.4% of annual vehicle miles traveled per household and 21.1% of annual vehicle trips per household (*6*). Therefore, the growth of online buying could have the potential to reduce traffic if it replaces physical shopping. However, if e-shopping induces new shopping trips, it is likely to generate more personal travel to existing transportation systems.

Although many studies have been conducted to model the adoption of e-shopping (*7*), little empirical work has been conducted to relate e-shopping to travel behavior, and the existing findings are limited to few geographical areas. By using 591 adult Internet users living in the Minneapolis–St. Paul metropolitan area of Minnesota, this study aims to provide further insight into the relationships between e-shopping and in-store shopping. A series of ordered probit models were developed to investigate their connections while controlling for a number of confounding factors. This research addresses the following central question: how, and to what extent, does e-shopping affect individuals' physical shopping?

LITERATURE REVIEW

Conceptual and empirical studies in the field of ICTs and transportation suggest that e-shopping may interact with travel behavior in four ways: substitution, complementarity, modification, and neutrality (*8–11*), although the connections can be even more complex (*4*). In the context of e-shopping, substitution denotes that a physical trip to traditional stores is replaced by an online transaction. Complementarity means that e-shopping generates new demands for trips to stores, either because of information obtained online or resulting from an online purchase. Modification denotes that e-shopping does not affect the amount of physical travel to stores but changes the characteristics of trips such as timing and chaining. Neutrality means that e-shopping is independent of traditional shopping. Among these effects, substitution carries important positive implications for congestion mitigation. Planners are eager to disentangle the complex connections between e-shopping and travel behavior and to measure the degree of substitution.

To assess the effects of e-shopping on individuals' shopping travel, why not just ask them? Several studies have found consistent substitution effects across many countries. Sim and Koi stated that 12% of 175 Singapore online buyers had reduced their travel to stores, although the remainder did not feel any influence (*12*). Tonn and Hemrick found that some Internet users in the Knoxville, Tennessee, metropolitan region reduced trips to stores, although a smaller percentage of users generated new trips (*13*). Weltevreden and van Rietbergen found that more than 20% of Dutch respondents

X. Cao and F. Douma, Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Avenue South, Minneapolis, MN 55455. F. Cleaveland, Minnesota Department of Transportation, 395 John Ireland Boulevard, St. Paul, MN 55155. Corresponding author: F. Douma, douma002@umn.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2157, Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 147–154.
DOI: 10.3141/2157-18

reported fewer trips to city center stores (*14*). Cairns et al. reported that 80% of 538 U.K. Internet users polled by British Telecom have saved at least one car trip because of online buying (*15*).

Some studies adopted a different questioning strategy and yielded different results. Wilson et al. asked respondents in three U.S. cities what they would have done if they had not found what they wanted in their last online purchase; 79% of respondents would have made a special trip to a store for that item, that is, online buying substituted for 79% of shopping trips (*16*). However, 21% of trips would not have happened; that is, these online purchases may not influence store shopping. They also found that 55.5% of respondents generated new trips because of product information obtained online. Corpuz and Peachman showed that 19% of respondents in Sydney, Australia, would have made a special trip if online buying were not available and that 14% would have not made the purchase (*17*). These two studies yielded mixed results of substitution, complementarity, and independence: online information may generate new trips to retail stores, some online purchases may replace shopping trips, and some may not affect shopping trips at all.

Going beyond descriptive questioning, a few studies applied multivariate models to reveal the relationships between e-shopping and travel behavior. By using the responses of 634 respondents to a shopping survey in Utrecht, the Netherlands, Farag et al. employed linear regression to explore the impact of online buying frequency on trip frequencies of daily in-store shopping and nondaily in-store shopping, controlling for sociodemographics, spatial attributes, Internet experience, and attitudes (*18*). They found that frequent online buyers were more likely to make shopping trips (for daily and nondaily products). They concluded that "the relationship between online buying and in-store shopping is not one of substitution but of complementarity" (*18*, p. 43).

Farag et al. applied path analysis to a sample of 826 respondents to a shopping survey in one urban municipality, Utrecht, and three surrounding suburban municipalities in the Netherlands (*19*). The model contained six endogenous variables: online buying frequency, online searching frequency, in-store shopping frequency, Internet use frequency, and two attitudes toward online buying and in-store shopping, respectively. The study assumed a unidirectional influence from e-shopping to in-store shopping. Farag et al. found that online searching frequency had a positive impact on in-store shopping frequency, although the influence of online buying frequency was insignificant. By using the same data set, Farag et al. augmented the path analysis to a structural equations model (SEM), in which traditional shopping and e-shopping were assumed to influence each other (*20*). For direct effects, they again found a positive association between online searching and in-store shopping. Further, people who shop often in stores tend to purchase often online (but not vice versa), controlled for shopping attitudes. Farag et al. concluded that a complementary relationship exists between online searching and in-store shopping, but online buying does not have a significant effect on in-store shopping.

Several studies have used activity-travel dairies to explore the implications of online buying for travel. On the basis of a supplementary survey, Zmud et al. classified respondents to the 1999 household travel survey in Sacramento, California, into two groups: Internet shopper and non-Internet shopper (*21*). They found that Internet shoppers had a higher overall trip rates than non-Internet shoppers and hence concluded that online buying may generate new trips. However, they did not control for any third-party variables such as income and hence did not rule out the possibility of spuriousness.

Ferrell used the 2000 San Francisco Bay Area Travel Survey (BATS 2000) to explore the relationships between home-based teleshopping (by Internet, catalog, and television) and travel behavior (*22*). He first developed two models to predict out-of-home shopping duration and nonshopping activity duration (with sociodemographics as explanatory variables) and then inserted predicted values into the equations of travel behavior. Ferrell found that teleshopping frequency has positive influences on shopping-trip frequency and trip-chaining frequency, but its influence on shopping-trip distance is insignificant. He concluded a complementary effect. Ferrell applied a SEM approach to the BATS 2000 to investigate the connections between teleshopping duration and shopping travel (*23*). He found that "people substitute home teleshopping time for shopping travel time and that teleshoppers take fewer shopping trips and travel shorter total distances for shopping purposes" (*23*, p. 212). Specifically, every 100-min teleshop is associated with a 5-min reduction in shopping travel time, with a 1-mi reduction in travel distance, and with 0.2-trip reduction in frequency. Ferrell did note that these two studies yielded contradictory relationships between teleshopping and shopping travel, although they used the same data. Ferrell attributed this critical difference to the limitations of the data set and different units of analysis in the two studies: household versus person.

It is becoming clear that the relationships between e-shopping and in-store shopping are more complex than they initially appeared: different survey methods yield different results, and different model specifications produce different results. In particular, studies that use the descriptive questioning approach tend to show both substitution and complementarity effects, and studies that use activity-travel diaries also produce mixed findings of substitution and complementarity. The studies that use shopping surveys found complementary and independent effects between online buying and traditional shopping, More important, the findings of existing multivariate analyses indicate only an association between e-shopping and personal travel but not causality.

METHODOLOGY

Data

The data for this study come from a shopping survey administered in the Minneapolis–St. Paul (Twin Cities) seven-county metropolitan area of Minnesota in 2008 and 2009. Because the purpose of this study is to investigate e-shopping via Internet, the study population is adult Internet users living in the Twin Cities area. To identify the sample frame, the authors hoped to acquire Internet subscriber data from local Internet providers. In the Twin Cities, Comcast and Qwest are the major providers of cable and DSL Internet connections, respectively. However, the data these companies can provide are inherently incomplete; Internet subscribers who use other service providers or connection types are excluded from their lists. Also, neither Comcast nor Quest readily provides subscriber data for research purposes. Therefore, households from chosen neighborhoods were sampled.

Because accurate Internet penetration data are not readily available, income was used as a proxy for Internet accessibility, and the sample was drawn from households living in certain census block groups with relatively high median household incomes. Census block groups were mapped by median household income in ArcGIS 9.2. Knowledge of the area was used to identify 16 sets of census block groups (totaling 78 block groups). Household median income in each of these block groups is well above the median household

income for the Twin Cities metropolitan statistical area ($54,304) with the median household income in most of the groups 50% to 100% higher.

Previous research showed that consumer adoption of e-shopping and the interaction between e-shopping and in-store shopping are likely to be influenced by retail accessibility (*14, 24*). Therefore, sample households were selected from three location types: urban neighborhoods, suburban neighborhoods, and exurban neighborhoods. These locations represent different levels of shopping accessibility. Because Internet penetration in exurban neighborhoods tends to be lower than that in the other two types of neighborhoods, exurban neighborhoods were oversampled by choosing four urban, four suburban, and eight exurban neighborhoods (16 sets of block groups) to achieve a relatively balanced sample of Internet users from all three location types. The geographical distribution of these neighborhoods is shown in Figure 1.

A random sample of 500 addresses in each of the 16 neighborhoods was purchased from AccuData Integrated Marketing, a commercial data provider (http://www.accudata.com). The survey was administered through Survey Monkey, a web-based survey provider, from December 2008 to January 2009. An invitation letter and two reminder postcards (1 and 4 weeks later) were mailed during the period. Five $100 gift cards were offered as an incentive to take the survey. After duplications and two nonadult participants were removed, the number of respondents totaled 591. The survey response rate was 7.85% (591 of 7,533, based on valid addresses only). However, this response rate is likely to be substantially lower than the true response rate because the population of this survey is Internet users, whereas many (although unknown) households in our sample frame are not Internet users.

Table 1 presents the number of respondents in each neighborhood, and Table 2 depicts selected demographic characteristics among

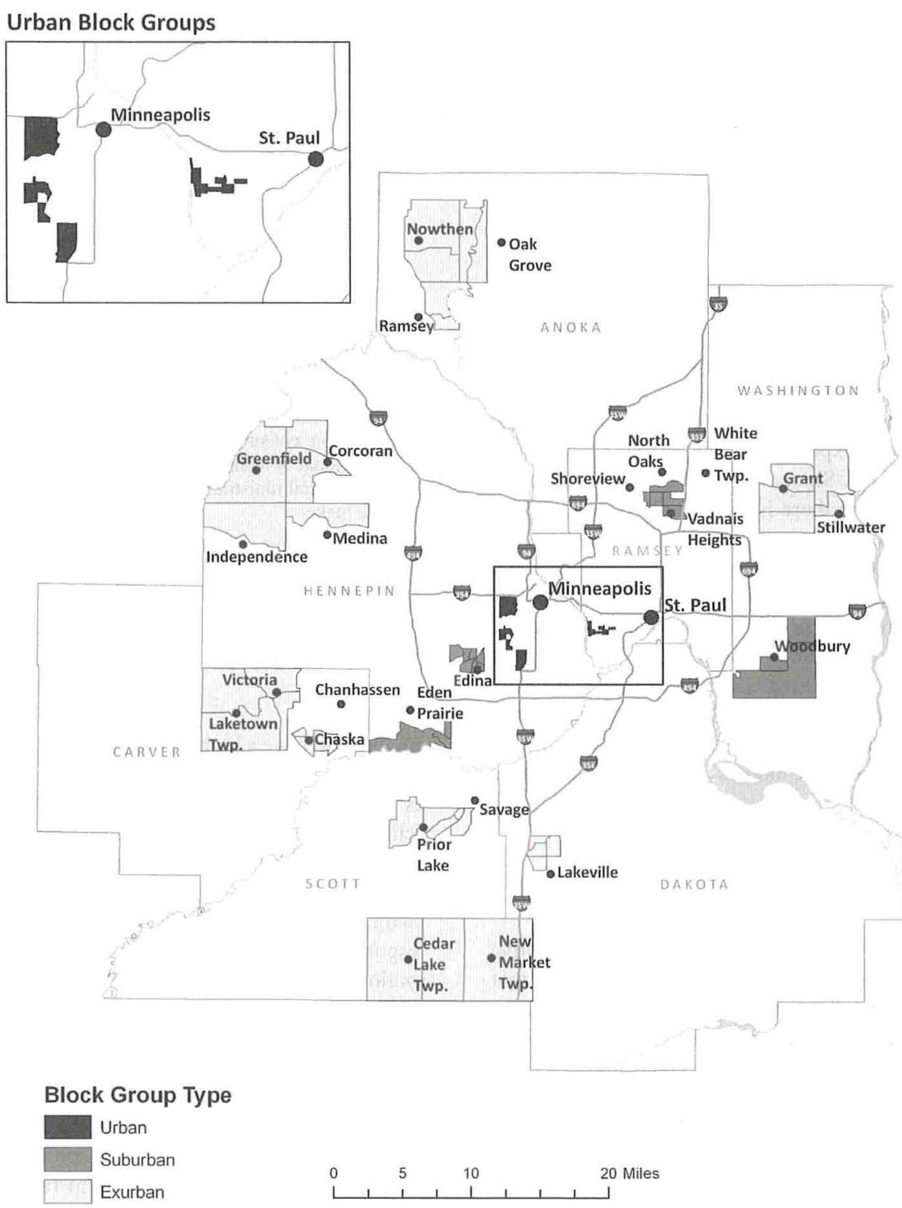


FIGURE 1 Target neighborhoods.

TABLE 1 Number of Respondents by Neighborhood

Neighborhood	Location	Respondents
Urban		
1	Linden Hills, Minneapolis	42
2	Lynhurst, Minneapolis	46
3	Macalester–Groveland, St. Paul	51
4	Cedar–Isles, Minneapolis	44
		183
Suburban		
5	Vadnais Heights	53
6	Eden Prairie	33
7	Woodbury	31
8	Edina	36
		153
Exurban		
9	Anoka	30
10	Rogers	26
11	Chaska	38
12	Waconia	39
13	Lakeville	33
14	Prior Lake	27
15	New Market	32
16	Stillwater	30
		255

NOTE: Response rates per neighborhood ($N = 591$).

TABLE 2 Sample Characteristics

Variable	Mean/Number	Percent
Age ($n = 566$)	48.39	
Number of vehicles per household ($n = 570$)	2.30	
Household size ($n = 569$)	2.81	
Gender ($n = 572$)		
Male	243	42.5
Female	329	57.5
Driver's license ($n = 570$)		
Yes	567	99.5
No	3	0.5
Credit card ($n = 570$)		
Yes	557	97.7
No	13	2.3
Household income ($n = 546$)		
Less than \$30,000	20	3.7
\$30,000–\$44,999	32	5.9
\$45,000–\$59,999	50	9.2
\$60,000–\$74,999	50	9.2
\$75,000–\$99,999	97	17.8
\$100,000–\$124,999	93	17.0
\$125,000–\$149,999	63	11.5
\$150,000 or more	141	25.8
Educational background ($n = 570$)		
High school or less	32	5.6
Some college or technical school	86	15.1
Two-year associate's degree	29	5.1
Four-year college or technical school degree	202	35.6
Some graduate school	56	9.9
Completed graduate school	163	28.7
Occupation ($n = 566$)		
Student	11	1.9
Sales	38	6.7
Service or repair	15	2.7
Manager or administrator	120	21.2
Production, construction, or crafts	23	4.1
Professional or technical	288	50.9
Clerical or administrative support	28	4.9
Homemaker	43	7.6
Employment status ($n = 568$)		
Work full-time	376	66.2
Work part-time	73	12.9
Not working	49	8.6
Retired	70	12.3

survey respondents. Generally, respondents are well-educated and well-paid workers with few limitations on their shopping behaviors such as low income or lack of a driver's license or credit card. The sample may not be a good representation of Internet users in the Twin Cities, especially when overdrawing high-income households. However, since the focus of the study is explaining the relationships of other variables to shopping behavior instead of on describing shopping behavior, this is not expected to materially affect the results of multivariate analysis (*25*). Since ample diversity of income is shown across the sample and there is thus ample opportunity to properly capture its effect on shopping behavior, the under- or over-representation of certain income categories is not serious. A similar principle holds regarding other sources of nonrepresentativeness, provided those sources are exogenous—such a situation is equivalent to exogenously stratified random sampling, and under those circumstances, coefficients of linear regression models are statistically unbiased (*26*). Thus, the sample does not claim perfect representativeness for univariate distributions, but the relationships among the variables of interest captured by the models (the focus of this study) are generalizable.

Variables

The rationale for variable development is as follows. Does a positive (or negative) association between shopping behaviors mean that e-shopping has a complementarity (or substitution) effect on in-store shopping? The answer is not straightforward, however. The positive association can be attributed to at least three mechanisms (*27*). First, online buying induces the need for additional related goods such as accessories (a direct causal influence). Second, time savings from online buying are reallocated to other shopping activities and travel (an indirect causal influence). Third, the association may result from factors antecedent to both shopping behaviors. Antecedent factors can result from various sources. For example, affluent individuals are likely to have a higher demand for online buying and traditional shopping than the poor; similarly, other sociodemographics (such as household size and age) can cause demand for both shopping activities in the same way. Shopping enjoyment may also increase the demand for shopping through multiple channels. Household shopping responsibility is another important, but mostly ignored, factor. Shopping is motivated by enjoyment but also by household responsibility. Failure to control for the responsibility may produce spurious relationships between online buying and traditional shopping at the individual level (*28*). Furthermore, shopping accessibility and location attributes may be third-party variables according to the efficiency hypothesis: people with a low shopping accessibility or in exurban neighborhoods may make fewer in-store purchases but conduct more online buying (*5*).

In this shopping survey, instruments were designed to account for the influences of the antecedent factors on both e-shopping and traditional in-store shopping. The variables in this study were classified into six groups: shopping behavior, shopping attitudes,

Internet experience, demographics, household responsibility, and shopping accessibility.

The variables of interest are respondent's e-shopping and in-store shopping behavior regarding nondaily purchases (products such as books, clothes, or electronics). These variables include frequency of store trips, frequency of Internet purchases, frequency of product information search via Internet, preferred alternative to Internet purchases, and a dummy indicating whether the respondent has made a store trip as a result of Internet browsing.

Variables for shopping attitudes were derived from a factor analysis. In the survey, respondents were asked to indicate whether they agree or disagree with 15 statements on a five-point Likert scale from "strongly disagree" to "strongly agree." The statements reflect fundamental shopping attitudes that are expected to influence respondents' shopping behavior. Since some of these statements measure similar dimensions of shopping attitudes and are highly correlated, a factor analysis was conducted to identify latent constructs of the attitudes. As shown in Table 3, four factor scores (price consciousness, time consciousness, shopping enjoyment, and impulsive shopper) were derived by using common factor analysis (called principle axis factoring in SPSS) with Varimax rotation.

The procedure for factor analysis is as follows. Before conducting factor analysis, a missing value analysis with an expectation maximum algorithm was applied to 15 statements because of 20 missing cells (out of $15 * 588 = 8{,}820$ cells) in the data. Because the purpose of the factor analysis is to identify latent constructs of shopping attitudes, common factor analysis (CFA) was used rather than principle components analysis (PCA). CFA tends to produce lower factor loading and percent variance explained by the factor solution than PCA. However, CFA is superior to PCA because the PCA loadings are more biased estimators than are the CFA loadings (*29*). A number of criteria were used in identifying the number of factors. In particular, both the eigenvalue-one rule and the scree rule point to the four-factor solution shown in the table. A five-factor solution was also tried. However, the statements underlying the fifth factor are a subset of the statements for the factor score of shopping enjoyment. Thus, the four-factor solution was chosen as the final results. Overall, 58.2% of total variation is explained by the factor scores using initial eigenvalues, and CFA explains 43.5% of total variation.

The category of Internet experience included variables describing the respondent's usage of the Internet: years of using Internet, frequency of personal Internet use, and Internet connection method (dial-up or broadband).

The demographic category included demographic characteristics that could influence the frequency of respondent's shopping activity, which consist of gender, age, possession of a driver's license, possession of a credit card, household income, number of household vehicles, household size, number of children, education background, student status, employment status, and any conditions preventing the respondent from driving, walking, or taking public transportation.

The household responsibility category included a single variable, the percent of household shopping activities for which the respondent is primarily responsible.

Shopping accessibility variables were derived from ArcGIS analysis. Retail accessibility was calculated for each of the 78 block groups included in this study. The U.S. Census Bureau's longitudinal employer–household dynamics data, which contains the number of retail jobs per block group in the Twin Cities area, were used for computing these variables. With number of retail jobs as a proxy for retail opportunity, five cumulative shopping accessibility variables were calculated for each block group: the number of retail jobs within a prespecified buffer (1, 2, 5, 10, and 20 mi). Because some block groups may partially fall within the prespecified buffer, the number of retail jobs in a block group was counted when the centroid of the block group fell within the buffer.

Analysis Approach

In the survey, in-store shopping frequency was measured on an ordinal scale ranging from 0 (less than once a month) to 3 (more than once a week). Specifically, 24.6% of respondents go to a store

TABLE 3 Rotated Factor Matrix of Factor Analysis

Statement	Shopping Enjoyment	Cost-Consciousness	Impulsive Shopper	Time Consciousness
I often make unplanned purchases.			0.662	
Before buying something, I generally take some time to think it over.		0.454	−0.352	
I generally stick to my shopping lists.			−0.552	
If I really want something, I'll often buy it even if it costs more than it should.		−0.354	0.428	
I generally compare prices before buying in order to get the best value.		0.751		
It's important to me to get the lowest prices when I buy things.		0.667		
I'm too busy to shop as often or as long as I'd like.				0.637
Being a smart shopper is worth the extra time it takes.		0.661		
Saving time when I shop is important to me.	−0.376			0.224
Under the right circumstances, shopping is fun.	0.608		0.239	0.227
Shopping is usually a chore for me.	−0.859			
For me, shopping can be an important leisure activity.	0.640		0.371	
Sometimes for me, shopping is mostly an excuse to get out of the house or workplace.	0.328		0.379	
Getting to where I usually shop is a hassle.	−0.372			0.221
Shopping at stores is too tiring to be enjoyable.	−0.765			

NOTE: The analysis is principle axis factoring with Varimax rotation, and loadings smaller than 0.2 were suppressed.

to purchase nondaily products for themselves and their households less than once a month (one respondent reported never); 47.9% go to a store one to three times per month; 20.3% go to a store once a week; and 7.2% go to a store more than once a week. Thus, ordered response techniques were used to model the behavior. In an ordered response model, individuals' behavior is observed in an ordinal scale ($Y = 0, 1, 2, \ldots, j, \ldots, J$), and it is assumed that an underlying latent continuous variable, Y^*, represents a respondent's propensity to store shopping (*30*). Y^* is expressed in the following form:

$$Y^* = \beta' X + e \tag{1}$$

where

X = vector of explanatory variables,
β = vector of parameters, and
e = unobserved error term.

The relationship between the latent Y^* and the observed Y is

$$Y = j \qquad \text{if } \mu_{j-1} < Y^* ! \mu_j,\ j = 0, 1, 2 \ldots, J \tag{2}$$

where the μ_j's are cut points or threshold parameters, defined as $\mu_{-1} = -\infty$, $\mu_J = +\infty$, and $\mu_{j-1} < \mu_j$ for all j. In the context of the ordered probit model, $e \sim N[0, \sigma_e^2]$, and one has the following probabilities:

$$P(Y = j) = P(\mu_{j-1} < Y^* \le \mu_j) = \Phi\left(\frac{\mu_j - \beta' X}{\sigma_e}\right) - \Phi\left(\frac{\mu_{j-1} - \beta' X}{\sigma_e}\right) \tag{3}$$

where ϕ denotes the standard normal cumulative distribution function. The coefficients are derived from maximum likelihood estimation of a sample.

The Veall–Zimmermann R^2 (R^2_{vz}) was chosen as a goodness-of-fit measure for the models because it is better than McFadden R^2 when the number of ordinal categories exceeds three (*31*). The R^2_{vz} is calculated as follows:

$$R^2_{vz} = \frac{[\mathrm{LL}(\beta) - \mathrm{LL}(C)][N - 2\mathrm{LL}(C)]}{-\mathrm{LL}(C)[N + 2\mathrm{LL}(\beta) - 2\mathrm{LL}(C)]} \tag{4}$$

where

LL(C) = log likelihood for the model where all coefficients but $J - 1$ thresholds are restricted to zero (equivalent to a constant-only model),
LL(β) = log likelihood at convergence (i.e., the final model), and
N = number of observations in the model.

RESULTS

Similar to that of Wilson et al. (*16*), this survey directly asked respondents what they would have done if their last Internet purchase had not been available via the Internet. Consistent with previous research, this question produces mixed results on the interactions between e-shopping and in-store shopping. In particular, about 29% of respondents would have made a special trip to the store to find this item or a similar item. This means that 29% of online purchases may substitute for in-store shopping. About 18% of respondents would have not made the purchase, which suggests that 18% of online purchases are likely to be induced demand. Further, 14% would have tried to purchase the item or a similar item through a catalogue or other media, a possible substitution for other teleshopping channels. Finally, 39% of respondents would have waited until their next trip to the store. Although the last three scenarios may not have generated additional trips to the store, that the respondent did purchase the item online likely resulted in some form of home delivery. In another question, 49.3% of respondents indicated that they have made a special trip to the store because of something they saw online, which provides evidence for the complementarity effect.

The SPSS 17.0 software package was used to develop five ordered probit models to explore the correlates of store shopping frequency. The first model illustrated the relationships between store shopping frequency and demographic components of the sample and those between store shopping behavior and Internet experience. The second incorporated data from variables regarding respondents' Internet use for online shopping and online searching. The third added neighborhood dummy variables and shopping accessibility variables. The fourth incorporated four factor scores of shopping attitudes. In the last specification, the variable indicating respondents' shopping responsibility in the household was incorporated in the model. The interest of this model is to show the relationship between in-store shopping and e-shopping behavior, with other variables being controls. When these models were developed, variables whose p-values were greater than 0.2 were dropped.

The results of these models are shown in Table 4. After some variables were dropped, Model 1 shows that household income is positively associated with store shopping frequency, although the association is not significant at the 0.05 level. The number of children (12 to 15 years old) has a positive association with store shopping frequency, indicating the larger demand induced by those children. Among students, full-time workers, and retirees, students were least likely to shop at a store frequently, followed by full-time workers, whereas retirees were most likely to shop at a store frequently. Further, respondents who had a broadband Internet connection tend to go shopping more frequently. Number of years of using the Internet is negatively associated with store shopping frequency, but the association is insignificant at the level of 0.05.

After demographics and Internet experience were controlled for, two more variables were found to be positively associated with store shopping frequency: how often respondents consulted the Internet for information about nondaily products and how often they actually bought those products over the Internet (Model 2). These two positive influences suggest that e-shopping (online searching and online buying) is likely to have a complementary effect on store shopping.

Model 3 further incorporated neighborhood location dummies and shopping accessibility variables. The suburban dummy is significant at the 0.05 level. In particular, suburban respondents were more likely to shop in stores. However, none of the remaining location variables are insignificant at the 0.2 level. In this model, e-shopping behavior variables remain statistically significant.

The fourth model added the shopping attitudes identified in the factor analysis. This model showed that three attitudes had positive effects on in-store shopping frequency, although the cost-consciousness variable is not significant even at the 0.1 level. As expected, those who intrinsically enjoy shopping as well as impulsive shoppers are more likely to shop in stores. In a comparison of the coefficients of online shopping frequency across the models, it was found that the coefficient in Model 4 is considerably larger than that in Model 3. In other words, once shopping attitudes were controlled for, the complementary connection between online shopping and store shopping becomes stronger. This is contrary to the expectation.

TABLE 4 Ordered Probit Models for In-Store Shopping Frequency

	Model 1		Model 2		Model 3		Model 4	
	Estimate	*p*-Value	Estimate	*p*-Value	Estimate	*p*-Value	Estimate	*p*-Value
Variable								
μ_0	−0.666	.059	−0.215	.558	−0.116	.753	−0.007	.984
μ_1	0.670	.057	1.137	.002	1.243	.001	1.423	.000
μ_2	1.540	.000	2.040	.000	2.154	.000	2.387	.000
Demographics and Internet Experience								
Household income	0.036	.146						
No. of children (12–15 years old)	0.179	.085	0.165	.107	0.170	.097	0.202	.051
Student	−0.626	.040	−0.581	.059	−0.551	.073	−0.703	.025
Full-time worker	−0.217	.065	−0.187	.104	−0.181	.116	−0.169	.146
Being retired	0.345	.038	0.377	.022	0.363	.027	0.356	.032
Years using Internet	−0.097	.177	−0.170	.017	−0.164	.020	−0.110	.127
Broadband connection	0.393	.021	0.280	.100	0.289	.089	0.315	.070
E-Shopping Behavior								
Frequency of product information search			0.191	.001	0.187	.001	0.138	.021
Frequency of online shopping			0.098	.011	0.103	.008	0.121	.002
Location: Suburb neighborhoods					0.243	.021	0.215	.044
Attitudes								
Shopping enjoyment							0.238	.000
Impulsive shopper							0.297	.000
Cost consciousness							0.078	.163
N	544		567		567		567	
Veall–Zimmermann *R*-squared	0.104		0.158		0.163		0.240	

If an antecedent variable affects two outcome variables in the same way, the association between the latter two variables tends to diminish once the antecedent variable is controlled for. By contrast, if the antecedent variable influences the two outcome variables in opposite ways, the association between outcomes becomes stronger when the influence of the antecedent variable on both outcomes is accounted for. This is equivalent to the concept of partial correlation, which measures the correlation between two variables while removing the effects of a third variable. To measure the correlation between online shopping frequency (X) and store shopping frequency (Y) while taking away the effects of shopping enjoyment (Z), the partial correlation between shopping frequencies can be derived by using the following equation:

$$r_{XY\cdot Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{\left(1-r_{XZ}^2\right)\left(1-r_{YZ}^2\right)}}$$

If the Pearson correlation between the frequencies r_{XY} is positive and both r_{XZ} and r_{YZ} are positive (or negative), the partial correlation $r_{XY\cdot Z}$ may be smaller than the Pearson correlation r_{XY}. If r_{XY} is positive, and one of r_{XZ} and r_{YZ} is positive and the other negative, the partial correlation $r_{XY\cdot Z}$ is larger than the Pearson correlation r_{XY}. Through a further exploration, it was found that the attitude of shopping enjoyment is positively associated with store shopping frequency but negatively associated with online shopping frequency. Therefore, it is reasonable that the coefficient of online shopping frequency in Model 4 is larger than that in Model 3. This result suggests that if shopping attitudes are not controlled for, there is a tendency to underestimate the complementarity between online shopping and store shopping.

The shopping responsibility variable was also included in the model. However, it is not significant at the 0.2 level. Thus, the model is not presented in the table. A bivariate correlation analysis confirmed that the variable is not significantly associated with either online shopping frequency or store shopping frequency. Therefore, this study failed to support the alternative hypothesis that shopping responsibility in the household is an antecedent factor of online and store shopping frequencies.

The Veall–Zimmermann R^2 increased from 0.104 in Model 1 to 0.240 in Model 4. The large increase results from e-shopping behavior and shopping attitudes. From these models, it is concluded that both online searching and online buying tend to have a complementary, rather than substitution, effect on in-store shopping trips. Although it appears that those who have used the Internet longest may have adopted behaviors that allowed them to reduce in-store shopping, increased access to the Internet in the form of a broadband connection, gathering information about a product via Internet, and online shopping allowed individuals to increase their shopping activities in stores.

CONCLUSIONS

This research sought to reveal the interactions between e-shopping and in-store shopping by using a sample of adult Internet users in the Minneapolis–St. Paul metropolitan area. Ordered probit models were

developed to account for the influences of a variety of confounding factors, such as shopping attitudes, shopping accessibility, shopping responsibility, and sociodemographics. The preliminary results show that online searching and online buying tend to have a complementary effect on in-store shopping, controlled for the confounding factors.

However, more-sophisticated modeling techniques could offer insight. The data showed that shopping enjoyment is negatively associated with online shopping frequency and positively associated with in-store shopping frequency. Shopping enjoyment was controlled for in the ordered probit models. However, the model did not establish the connection between shopping enjoyment and online shopping frequency. This type of disconnection may be true for other variables in the model. Therefore, a structural equations model may be useful to capture various interactions among variables.

ACKNOWLEDGMENT

This research was funded by the Tech Plan Program of the ITS Institute at the University of Minnesota.

REFERENCES

1. *E-Commerce Sales Rise 19% Last Year in U.S.* http://www.InternetRetailer.com. 2008.
2. Farag, S. *E-Shopping and Its Interactions with In-Store Shopping.* Utrecht University, Netherlands, 2006.
3. Cao, X., and P. Mokhtarian. *The Intended and Actual Adoption of Online Purchasing: A Brief Review of Recent Literature.* UCD-ITS-RR-05-07. Institute of Transportation Studies, University of California, Davis, 2005.
4. Mokhtarian, P. L. A Conceptual Analysis of the Transportation Impacts of B2C e-Commerce. *Transportation,* Vol. 31, No. 3, 2004, pp. 257–284.
5. Anderson, W. P., L. Chatterjee, and T. R. Lakshmanan. E-Commerce, Transportation, and Economic Geography. *Growth and Change,* Vol. 34, No. 4, 2003, pp. 415–432.
6. Hu, P. S., and T. R. Reuscher. *Summary of Travel Trends: 2001 National Household Travel Survey.* FHWA, U.S. Department of Transportation, 2004.
7. Chang, M. K., W. M. Cheung, V. S. Lai. Literature Derived Reference Models for the Adoption of Online Shopping. *Information and Management,* Vol. 42, No. 4, 2005, pp. 543–559.
8. Choo, S., T. Lee, and P. L. Mokhtarian. Do Transportation and Communications Tend to Be Substitutes, Complements, or Neither?: U.S. Consumer Expenditures Perspective, 1984–2002. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2010,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 121–132.
9. Mokhtarian, P. L. A Typology of Relationships Between Telecommunications and Transportation. *Transportation Research Part A,* Vol. 24, No. 2, 1990, pp. 231–242.
10. Pendyala, R. M., K. G. Goulias, and R. Kitamura. Impact of Telecommuting on Spatial and Temporal Patterns of Household Travel. *Transportation,* Vol. 18, No. 4, 1991, pp. 383–409.
11. Salomon, I. Telecommunications and Travel Relationships: A Review. *Transportation Research Part A,* Vol. 20, No. 2, 1986, pp. 223–238.
12. Sim, L. L., and S. M. Koi. Singapore's Internet Shoppers and Their Impact on Traditional Shopping Patterns. *Journal of Retailing and Consumer Services,* Vol. 9, No. 2, 2002, pp. 115–124.
13. Tonn, B. E., and A. Hemrick. Impacts of the Use of E-Mail and the Internet on Personal Trip-Making Behavior. *Social Science Computer Review,* Vol. 22, No. 2, 2004, pp. 270–280.
14. Weltevreden, J. W. J., and T. Van Rietbergen. E-Shopping Versus City Centre Shopping: The Role of Perceived City Centre Attractiveness. *Tijdschrift Voor Economische En Sociale Geografie,* Vol. 98, 2007, pp. 68–85.
15. Cairns, S., L. Sloman, C. Newson, J. Anable, A. Kirkbride, and P. Goodwin. *Smarter Choices: Changing the Way We Travel.* Department for Transport, London, 2004.
16. Wilson, R. D., K. J. Krizek, and S. L. Handy. Trends in Out-of-Home and At-Home Activities: Evidence from Repeat Cross-Sectional Surveys. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2014,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 76–84.
17. Corpuz, G., and J. Peachman. Measuring the Impacts of Internet Usage on Travel Behaviour in the Sydney Household Travel Survey. Presented at Australasian Transport Research Forum Conference, Wellington, New Zealand, 2003.
18. Farag, S., K. J. Krizek, and M. Dijst. E-Shopping and Its Relationship with In-Store Shopping: Empirical Evidence from the Netherlands and the USA. *Transport Reviews,* Vol. 26, 2006, pp. 43–61.
19. Farag, S., T. Schwanen, and M. Dijst. Empirical Investigation of Online Searching and Buying and Their Relationship to Shopping Trips. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1926,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 242–251.
20. Farag, S., T. Schwanen, M. Dijst, and J. Faber. Shopping Online and/or In-Store? A Structural Equation Model of the Relationships Between E-Shopping and In-Store Shopping. *Transportation Research Part A,* Vol. 41, No. 2, 2007, pp. 125–141.
21. Zmud, J., S. G. Bricka, and J. Casas. Impact of Shopping via Internet on Travel for Shopping Purposes. Presented at 80th Annual Meeting of the Transportation Research Board, Washington, D.C., 2001.
22. Ferrell, C. Home-Based Teleshoppers and Shopping Travel: Do Teleshoppers Travel Less? In *Transportation Research Record: Journal of the Transportation Research Board, No. 1894,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 241–248.
23. Ferrell, C. E. Home-Based Teleshopping and Shopping Travel: Where Do People Find the Time? In *Transportation Research Record: Journal of the Transportation Research Board, No. 1926,* Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 212–223.
24. Ren, F., and M. P. Kwan. The Impact of Geographic Context on E-Shopping Behavior. *Environment and Planning B: Planning and Design,* Vol. 36, No. 2, 2009, pp. 262–278.
25. Babbie, E. R. *The Practice of Social Research,* 11th ed, Thomson Wadsworth, Belmont, Calif., 2007.
26. Hausman, J., and D. Wise. Stratification on Endogenous Variables and Estimation: The Gary Income Maintenance Experiment. In *Structural Analysis of Discrete Data with Econometric Applications* (C. Manski and D. McFadden, eds.), MIT Press, Cambridge, Mass., 1981, pp. 365–391.
27. Mokhtarian, P. L., and G. Circella. The role of Social Factors in Store and Internet Purchase Frequencies of Clothing/Shoes. Presented at International Workshop on Frontiers in Transportation: Social Interactions, Amsterdam, Netherlands, 2007.
28. Cao, X. E-Shopping, Spatial Attributes, and Personal Travel: A Review of Empirical Studies. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2135,* Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 160–169.
29. Widaman, K. F. Common Factor Analysis Versus Principal Component Analysis: Differential Bias in Representing Model Parameters? *Multivariate Behavioral Research,* Vol. 28, 1993, pp. 263–311.
30. Daykin, A. R., and P. G. Moffatt. Analyzing Ordered Responses: A Review of the Ordered Probit Model. *Understanding Statistics,* Vol. 1, 2002, pp. 157–166.
31. Veall, M. R., and K. F. Zimmermann. Pseudo-R2 Measures for Some Common Limited Dependent Variable Models. Presented at Sonderforschungsbereich 386, 1996.

The Telecommunications and Travel Behavior Committee peer-reviewed this paper.

TRANSPORTATION RESEARCH RECORD: JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Peer Review Process

The ***Transportation Research Record: Journal of the Transportation Research Board*** publishes approximately 25% of the more than 3,000 papers that are peer reviewed each year. The mission of the Transportation Research Board (TRB) is to disseminate research results to the transportation community. The Record series contains applied and theoretical research results as well as papers on research implementation.

The TRB peer review process for the publication of papers allows a minimum of 30 days for initial review and 60 days for rereview, to ensure that only the highest-quality papers are published. At least three reviews must support a committee's recommendation for publication. The process also allows for scholarly discussion of any paper scheduled for publication, along with an author-prepared closure.

The basic elements of the rigorous peer review of papers submitted to TRB for publication are described below.

Paper Submittal: June 1–August 1

Papers may be submitted to TRB at any time. However, most authors use the TRB web-based electronic submission process available between June 1 and August 1, for publication in the following year's Record series.

Initial Review: August 15–November 15

TRB staff assigns each paper by technical content to a committee that administers the peer review. The committee chair assigns at least three knowledgeable reviewers to each paper. The initial review is completed by mid-September.

By October 1, committee chairs make a preliminary recommendation, placing each paper in one of the following categories:

1. Publish as submitted or with minor revisions,
2. Publish pending author changes and rereview, or
3. Reject for publication.

By late October, TRB communicates the results of the initial review to the corresponding author indicated on the paper submission form. Corresponding authors communicate the information to coauthors. Authors of papers in Category 2 (above) must submit a revised version addressing all reviewer comments and must include a cover letter explaining how the concerns have been addressed.

Rereview: November 20–January 25

The committee chair reviews revised papers in Category 1 (above) to ensure that the changes are made and sends the Category 2 revised papers to the initial reviewers for rereview. After rereview, the chairs make the final recommendation on papers in Categories 1 and 2. If the paper has been revised to the committee's satisfaction, the chair will recommend publication. The chair communicates the results of the rereview to the authors.

Discussions and Closures: February 1–May 15

Discussions may be submitted for papers that will be published. TRB policy is to publish the paper, the discussion, and the author's closure in the same Record.

Many papers considered for publication in the *Transportation Research Record* are also considered for presentation at TRB meetings. Individuals interested in submitting a discussion of any paper presented at a TRB meeting must notify TRB no later than February 1. If the paper has been recommended for publication in the *Transportation Research Record*, the discussion must be submitted to TRB no later than April 15. A copy of this communication is sent to the author and the committee chair.

The committee chair reviews the discussion for appropriateness and asks the author to prepare a closure to be submitted to TRB by May 15. The committee chair reviews the closure for appropriateness. After the committee chair approves both discussion and closure, the paper, the discussion, and the closure are included for publication together in the same Record.

Final Manuscript Submittal: March 15

In early February, TRB requests a final manuscript for publication—to be submitted by March 15—or informs the author that the paper has not been accepted for publication. All accepted papers are published by December 31.

Paper Awards: April to January

The TRB Executive Committee has authorized annual awards sponsored by Groups in the Technical Activities Division for outstanding published papers:

- Charley V. Wootan Award (Policy and Organization Group);
- Pyke Johnson Award (Planning and Environment Group);
- K. B. Woods Award (Design and Construction Group);
- Patricia F. Waller Award (Safety and System Users Group);
- D. Grant Mickle Award (Operations and Preservation Group); and
- John C. Vance Award (Legal Resources Group).

Other Groups also may nominate published papers for any of the awards above. In addition, each Group may present a Fred Burggraf Award to authors 35 years of age or younger.

Peer reviewers are asked to identify papers worthy of award consideration. Each Group reviews all papers nominated for awards and makes a recommendation to TRB by September 1. TRB notifies winners of the awards, which are presented at the following TRB Annual Meeting.

Transportation Research Board
www.TRB.org

TRANSPORTATION RESEARCH BOARD
500 Fifth Street, NW
Washington, DC 20001

ADDRESS SERVICE REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 8970

THE NATIONAL ACADEMIES™

Advisers to the Nation on Science, Engineering, and Medicine

The nation turns to the National Academies—National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council—for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org

ISBN 978-0-309-14283-0
90000
9 780309 142830